

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
19 September 2002 (19.09.2002)

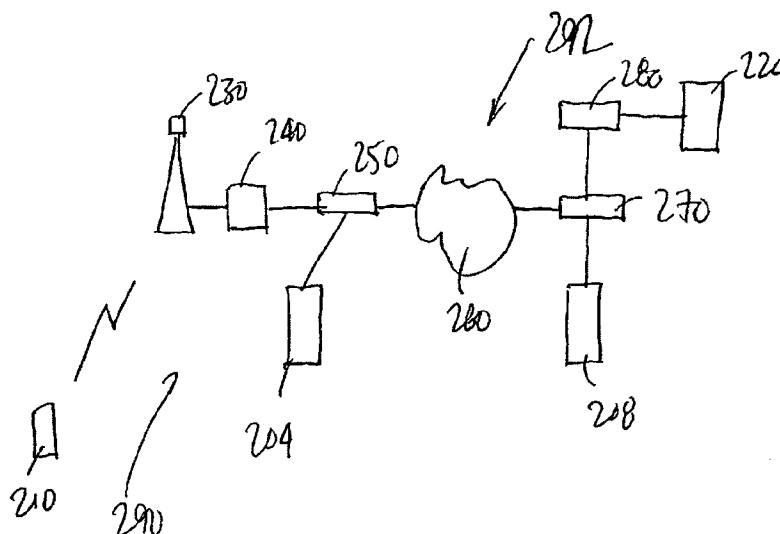
PCT

(10) International Publication Number  
**WO 02/073443 A1**

- (51) International Patent Classification<sup>7</sup>: **G06F 15/16, H03M 5/00**
- (21) International Application Number: PCT/US02/08218
- (22) International Filing Date: 13 March 2002 (13.03.2002)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:  
60/275,584 13 March 2001 (13.03.2001) US  
10/099,901 12 March 2002 (12.03.2002) US
- (71) Applicant: **DILITHIUM NETWORKS, INC.** [US/US];  
700 Larkspur Landing Circle, Suite 263, Larkspur, CA  
94939 (US).
- (72) Inventor: **JABRI, Marwan, Anwar**; Unit 44, 267 Cast-  
ereagh Street, Sydney, New South Wales 2000 (AU).
- (74) Agents: **OGAWA, Richard, T.** et al.; Townsend and  
Townsend and Crew LLP, 2 Embarcadero Center, 8th  
Floor, San Francisco, CA 94111 (US).
- (81) Designated States (*national*): AE, AG, AL, AM, AT, AU,  
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,  
CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH,  
GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,  
LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW,  
MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG,  
SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU,  
ZA, ZM, ZW.
- (84) Designated States (*regional*): ARIPO patent (GH, GM,  
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),  
Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),  
European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR,  
GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent  
(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR,  
NE, SN, TD, TG).
- Published:  
— with international search report

[Continued on next page]

(54) Title: METHOD AND SYSTEM FOR TRANSCODING VIDEO AND SPEECH SIGNALS



(57) Abstract: A system and method for transferring multimedia information from a source location (210) to a destination location (220) through one or more networks (260, 290, 292), which may be different. The system has a source output which provides a first stream of information in a first format. The system also has a destination input which receives a second stream of information in a second format. A proxy transcoder server ("PTS") (204, 208) is coupled between the source output and the destination input. The PTS has a transcoding module transcoding data. The PTS has a transcoding module transcoding data. The PTS also has a capability module identifying a first capability of the source output and a second capability of the destination input, and selecting a transcoding process based upon the first capability and the second capability.

WO 02/073443 A1

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2004-534424

(P2004-534424A)

(43) 公表日 平成16年11月11日(2004. 11. 11)

(51) Int. Cl.<sup>7</sup>

H04L 12/66

F I

H04L 12/66

E

テーマコード (参考)

5K030

審査請求 未請求 予備審査請求 有 (全 126 頁)

(21) 出願番号 特願2002-572034 (P2002-572034)  
 (86) (22) 出願日 平成14年3月13日 (2002. 3. 13)  
 (85) 翻訳文提出日 平成15年9月16日 (2003. 9. 16)  
 (86) 国際出願番号 PCT/US2002/008218  
 (87) 国際公開番号 WO2002/073443  
 (87) 国際公開日 平成14年9月19日 (2002. 9. 19)  
 (31) 優先権主張番号 60/275, 584  
 (32) 優先日 平成13年3月13日 (2001. 3. 13)  
 (33) 優先権主張国 米国 (US)  
 (31) 優先権主張番号 10/099, 901  
 (32) 優先日 平成14年3月12日 (2002. 3. 12)  
 (33) 優先権主張国 米国 (US)

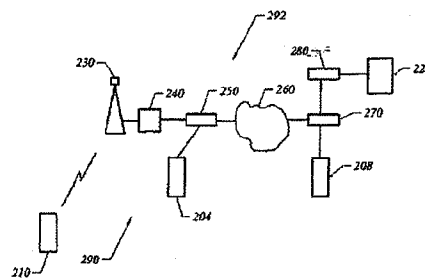
(71) 出願人 503336372  
 マクチャイナ ビーティーワイ リミテッ  
 ド  
 オーストラリア国 ニュー サウス ウェ  
 ールズ州 ブロードウェイ スメイル ス  
 トリート レベル 7  
 (74) 代理人 100102978  
 弁理士 清水 初志  
 (74) 代理人 100108774  
 弁理士 橋本 一憲  
 (72) 発明者 ジャブリ マルワン アンワー  
 オーストラリア国 ニュー サウス ウェ  
 ールズ州 シドニー キャステリーグ ス  
 トリート 267 ユニット 44

最終頁に続く

(54) 【発明の名称】 ビデオ信号および音声信号をコード変換する方法およびシステム

## (57) 【要約】

マルチメディア情報を、互いに異なっていてよい1つまたは複数のネットワーク(260、290、292)を通じて、送信元位置(210)から宛先位置(220)へ転送するシステムおよび方法。システムは、第1のフォーマットの第1の情報ストリームを提供する送信元出力を有する。システムは、第2のフォーマットの第2の情報ストリームを受信する宛先入力も有する。送信出力と宛先入力との間に代理コード変換サーバ(「PTS」)が結合されている。PTSは、データをコード変換するコード変換モジュールを有する。PTSは、送信元出力の第1の機能および宛先入力の第2の機能を識別し、かつ第1の機能および第2の機能に基づいてコード変換プロセスを選択する、機能モジュールも有する。



**【特許請求の範囲】****【請求項 1】**

マルチメディア情報を、1つまたは複数のネットワークを通じて送信元位置から宛先位置へ転送するためのシステムであって：

複数の送信元のうちの1つにおいて第1の情報ストリームを提供する送信元出力；

複数の宛先機能のうちの1つにおいて第2の情報ストリームを受信する宛先入力；

送信元出力と宛先入力との間に結合された代理コード変換サーバ（「PTS」）

を含み、ここで代理コード変換サーバは、

送信元出力の送信元機能を識別するように適合し、かつ宛先入力の宛先機能を識別するように適合した機能モジュール；

送信元機能のうちの1つの機能および宛先機能のうちの1つの機能に基づいて、コード変換プロセスを選択するように適合した選択モジュール；

選択されたコード変換プロセスを用いて、第1の情報ストリームを処理するように適合したコード変換モジュール；

を含むシステム。

**【請求項 2】**

1つまたは複数のトランスポート・ネットワークは、インターネット、移動ネットワーク、広域ネットワーク、ローカル・エリア・ネットワーク、PTSN、ISDN、およびSONETを含む群から選択される、請求項1記載のシステム。

**【請求項 3】**

送信元出力および宛先入力の少なくとも一方は、リモート装置のものである、請求項1記載のシステム。

**【請求項 4】**

機能モジュールは、リモート装置に記憶されている情報に基づき、ユーザのサービス・プロバイダのネットワーク・データベースに記憶されているユーザ加入情報に基づき、サービス・プロバイダによって交換されるか、または事前に設定されるストリーム内の、帯域内情報コマンドおよび制御に基づき、リモート装置の出力と入力の少なくとも一方を識別する、請求項3記載のシステム。

**【請求項 5】**

機能モジュールによって選択されたコード変換プロセスは、データを第1のビットストリーム・プロトコル・モードから第2のビットストリーム・プロトコル・モードにコード変換する、請求項1記載のシステム。

**【請求項 6】**

PTSは、PTSによって生成されるデータ・レートを調整するレート制御モジュールをさらに含む、請求項1記載のシステム。

**【請求項 7】**

レート制御モジュールは、ネットワーク・ホストもしくはネットワーク・アクセス・プロバイダ、または内部PTS機構から得たネットワーク輻輳情報、帯域幅情報、品質情報に基づいて「往復」時間情報を算出することによって、ネットワーク・ステータス情報を検出する、請求項6記載のシステム。

**【請求項 8】**

「往復」時間情報は、送信元位置または宛先位置のいずれかに「ピング」パケットを送信することによって測定することができる、請求項7記載のシステム。

**【請求項 9】**

レート制御モジュールは、帯域内情報を用いることによってネットワーク・ステータス情報を検出する、請求項6記載のシステム。

**【請求項 10】**

レート制御モジュールは、コード変換パラメータを変更することによってデータ・レートを調整する、請求項6記載のシステム。

**【請求項 11】**

レート制御モジュールは、ネットワーク機器に、PTSが対処中のデータに他のデータよりも高い優先順位を与えるよう命令することによってデータ・レートを調整する、請求項6記載のシステム。

【請求項12】

機能のフォーマットは、ITU、IETE、およびWAPを含む群から選択される、請求項1記載のシステム。

【請求項13】

1つまたは複数のネットワークは、各々が特定の標準用に構成されている、複数の異なるネットワークから選択される、請求項1記載のシステム。

【請求項14】

PTSは、送信元出力のネットワーク・アドレスおよび宛先入力 of ネットワーク・アドレスを決定する、ネットワーク・アドレス指定モジュールをさらに含む、請求項1記載のシステム。

【請求項15】

PTSは、2つまたはそれ以上のオーディオ・ストリームに関連するビットストリームを組み合わせ、組み合わせられたビットストリームを宛先入力に再送する、メディア混合プロセスをさらに含む、請求項1記載のシステム。

【請求項16】

PTSは、知的所有権に関する情報を管理かつ処理する、知的所有権管理モジュールをさらに含む、請求項1記載のシステム。

【請求項17】

PTSは、データを暗号化かつ復号する、暗号化および復号プロセスをさらに含む、請求項1記載のシステム。

【請求項18】

レート制御モジュールは、データ・レートを動的にかつリアルタイムに調整する、請求項6記載のシステム。

【請求項19】

コード変換モジュールは、送信元出力用の様々な種類の機能と宛先入力用の様々な種類の機能との間でコード変換を行うようにプログラム可能である、請求項1記載のシステム。

【請求項20】

マルチメディア情報を、1つまたは複数のネットワークを通じて送信元位置から宛先位置へ転送するためのシステムであって：

第1のネットワークに結合され、第1の情報ストリームを提供する、複数の送信元機能からの第1のフォーマットでの送信元出力；

第2のネットワークに結合され、第2の情報ストリームを受信する、複数の宛先機能から第2のフォーマットで受信される宛先入力；

送信元出力と宛先入力との間に結合された代理コード変換サーバ（「PTS」）

を含み、ここで代理コード変換サーバは、

送信元出力の第1のフォーマットを識別するように適合し、かつ宛先入力の第2のフォーマットを識別するように適合した機能プロセス；

1からN（Nは1より大きい整数）まで番号付けされた複数のコード変換モジュールを含み、機能に関連する第1のフォーマットおよび第2の機能に関連する第2のフォーマットに基づいて、1つのコード変換プロセスを選択するように適合した、機能プロセスに結合されたコード変換プロセス；

第1のネットワークからネットワーク・ステータス情報を受信するように適合し、ネットワーク・ステータス情報に基づいて情報ストリームのステータスを調整するように適合した、コード変換プロセスに結合されたビット・レート制御プロセス

を含むシステム。

【請求項21】

ステータス情報はピングを含む、請求項20記載のシステム。

**【請求項 2 2】**

ステータスは停止ステータスである、請求項20記載のシステム。

**【請求項 2 3】**

ステータスは優先順位付けステータスである、請求項20記載のシステム。

**【請求項 2 4】**

ステータスは、下位ビット・レート・コードを選択することによってビット・レートを調整するものである、請求項20記載のシステム。

**【請求項 2 5】**

情報ストリームを処理する方法であって、

情報ストリーム用の複数の送信元機能から送信元機能を識別する段階； 10

複数の宛先機能から宛先機能を識別する段階；

識別された送信元機能および識別された宛先機能に基づいて、ライブラリ内の複数のコード変換プロセスからコード変換プロセスを選択する段階；

識別された送信元機能および識別された宛先機能が異なる場合に、選択されたコード変換プロセスを用いて情報ストリームを処理する段階；

識別された送信元機能および識別された宛先機能が一致する場合に、1つのコード変換プロセスとは無関係に情報ストリームを送信元から宛先に転送する段階

を含む方法。

**【請求項 2 6】**

選択されるコード変換プロセスは、経験的な情報によって提供される、請求項25記載の方法。 20

**【請求項 2 7】**

ライブラリは、少なくとも複数の送信元機能および複数の宛先機能を、第2の次元に有する参照テーブルである、請求項25記載の方法。

**【発明の詳細な説明】****【技術分野】****【0 0 0 1】**

関連出願の相互参照

本出願は、2001年5月13日に出願され、共通の譲渡人を有し、参照として本明細書に組み入れられる米国仮特許出願第60/275,584号に対する優先権を主張する。 30

**【0 0 0 2】**

連邦が後援する研究または開発の下でなされた発明の権利に関する声明

該当なし

**【0 0 0 3】**

コンパクト・ディスク上で提出される「配列表」、表、コンピュータ・プログラム表付属文書への参照

該当なし

**【背景技術】****【0 0 0 4】**

本発明は通信の分野に関する。特に、本発明は、ビデオ信号およびオーディオ信号をコード変換する方法および装置に関する。さらに、本発明は、少なくとも1つのプロセスが情報をコード変換するために選択される、複数のコード変換プロセスを有する代理コード変換サーバを用いて、情報（たとえば、ビデオ、音声、データ）を第1のフォーマットから宛先フォーマットにコード変換する方法およびシステムを提供する。単に例示のため、本発明は広域電気通信ネットワークに適用されるが、本発明をインターネット、移動ネットワーク、ローカル・エリア・ネットワーク、PTSN、ISDN、SONET、DWDMなどのトランスポート・ネットワーク上で多数の異なる種類のマルチメディア・プロトコルを介して適用することもできることが認識される。 40

**【0 0 0 5】**

過去数年間のうちに電気通信技術は大幅に向上している。固定交換ネットワーク、パケッ 50

ト・ベースのネットワーク、移動ネットワークのような多数の異なる種類のネットワークが配備されている。「インターネット」と呼ばれる最も広く知られている広域ネットワークによって、ネットワーク化は世界中の多数の人に普及している。インターネットなどの広域ネットワークが盛んに利用されるようになったため、電子メール、ビデオ電話、ビデオ・ストリーミング、電子商取引のような多数の新しいオンライン・サービスが生まれている。最初はインターネットにコンピュータが接続されたが、携帯電話、パーソナル・デジタル・アシスタンス、ラップトップ・コンピュータのような他の装置も接続されている。したがって、現在、多数の異なる種類の装置が様々なネットワーク上の多数の異なる種類のサービスにアクセスしている。

#### 【0006】

10

前述の装置を互いに接続する様々なネットワーク要素がネットワークを構成している。このような装置は、データの伝送、および送信側ネットワークのプロトコルから受信側ネットワークによって使用されているプロトコルへのメッセージの変換に対処する、ゲートウェイおよび交換機によって接続されることが多い。ゲートウェイおよび交換機は、アナログ音声メッセージを、ITU標準であるG.711およびG.723.1を含むデジタル・フォーマットに変換する。ゲートウェイは通常、IP上での音声の送信と同様の方法で、変換されたメッセージを送信する。G.711は、音声コーデック用のITU標準であり、A-Law PCM法またはμ-Law PCM法を用いて64Kbpsのオーディオ信号を与える。G.723.1は、Plain Old Telephone Systemsおよび狭帯域インターネット接続を含む、狭帯域ネットワークに最適化された音声コーデック用のITU標準である。この標準は、LD-CELP法を使用し、5.3Kbpsまたは6.3Kbpsのオーディオ信号を提供する。用途に応じて他の多数の標準を用いることができる。

#### 【0007】

図1には単なる一例として従来のシステム100が示されている。この図は、一例に過ぎず、例示のためのみに提供されている。メッセージは、無線ネットワークに結合された移動装置105から発信される。メッセージは移動装置から無線ネットワークを通じて基地局110に送信される。基地局は、ゲートウェイ120に結合されたサービス局115に結合されている。基地局は、移動装置105から無線メッセージを受信し、メッセージをコード変換せずにデジタル・フォーマットに変換し、それをサービス局に送信する。再フォーマットされたメッセージはその後、ゲートウェイに送信され、ゲートウェイは、インターネット125を通じ、かつ様々なネットワーク要素を通じて、このメッセージをその宛先、すなわちユーザに送信する。このような要素にはゲートウェイ130、サーバ135などが含まれる。

30

#### 【0008】

1つまたは複数のゲートウェイは、テレビ会議信号をあるデジタル・フォーマットから他のデジタル・フォーマットへ、たとえばH.320からH.323に変換し、変換された信号をインターネット上で送信する。H.320は、デジタル回線上のテレビ会議用のITU標準であり、H.261ビデオ圧縮法を使用することにより、H.320に準拠したテレビ会議システムおよびデスクトップ・システムが、ISDN、交換デジタル回線、および専用回線を介して互いに通信するのを可能にする。H.323は、LANおよびインターネット上のリアルタイム対話型音声およびビデオ会議用のITU標準である。H.323は、IP電話に広く使用されており、音声、ビデオ、およびデータの任意の組合せをトランスポートするのを可能にする。H.323は、H.261およびH.263を含むいくつかのビデオ・コーデック、ならびにG.711およびG.723.1を含むオーディオ・コーデックを特定している。残念なことに、オーディオ標準およびビデオ標準は、H.320、H.323、G.711、およびG.723.1をはるかに超えて発達してきた。すなわち、様々な標準が存在するため、それらの間でメッセージを伝達することが困難になっている。さらに、このような標準間の通信のために、時間がかかり効率の悪い複雑な変換技術が増えている。したがって、情報またはコード変換を様々なフォーマット間でリアルタイムに変換する効率的な方法が必要である。H.320やH.324のような回線交換システムもあれば（データがビットの連続的なストリームとして送信される）、パケット・ベースのシステムもあるので、回線ベースのシステムとパケット・ベースのシステムを接続するには、回線ベースのビットをパケットに逆多重化し（回線からパケット）、かつその逆も同様に行う

40

50

(パケットから回線) 必要がある。H.320、H.323、H.324、3GPP-324M、SIP、およびSDPのような様々なシステム・プロトコルが(接続をセットアップし、端末機能を交換するために)異なるシグナリング方法を使用することに留意されたい。これらのシステムを相互接続するには、それぞれの異なるプロトコルを使用する端末が何を実行できるのかを各端末が理解できるように、トランスシグナリングおよび端末機能の変換が必要である。

【0009】

上記のことから、情報を送信元から宛先に転送する改良方法が極めて望ましいことが分かる。

【発明の開示】

【0010】

本発明によれば、電気通信分野におけるコード変換のための改良技術が提供される。特に、本発明は、ビデオ信号および音声信号をコード変換する方法および装置に関する。さらに、本発明は、少なくとも1つのプロセスが情報をコード変換するために選択される、複数のコード変換プロセスを有する代理コード変換サーバを用いて、情報(たとえば、ビデオ、音声、データ)を第1のフォーマットから宛先フォーマットにコード変換する方法およびシステムを提供する。単なる例示のため、本発明は広域電気通信ネットワークに適用されるが、本発明を、インターネット、移動ネットワーク、ローカル・エリア・ネットワーク、PTSN、ISDN、SONET、DWDMなどのトランスポート・ネットワーク上で多数の異なる種類のマルチメディア・プロトコルを介して適用することもできることが認識される。

【0011】

特定の態様では、本発明は、マルチメディア情報を、互いに異なっていてよい1つまたは複数のネットワークを通じて、送信元位置から宛先位置へ転送するシステムを提供する。システムは、第1のフォーマットの第1の情報ストリームを提供する送信元出力を有する。システムは、第2のフォーマットの第2の情報ストリームを受信する宛先入力も有する。送信出力と宛先入力との間に代理コード変換サーバ(「PTS」)が結合されている。PTSは、データをコード変換するコード変換モジュールを有する。PTSは、送信元出力の第1の機能および宛先入力の第2の機能を識別し、第1の機能および第2の機能に基づいてコード変換プロセスを選択する機能モジュールも有する。好ましくは、選択は機能モード選択を用いて提供される。

【0012】

他の特定の態様では、本発明は、マルチメディア情報を、互いに異なっていてよい1つまたは複数のネットワークを通じて、送信元位置から宛先位置へ転送するシステムを提供する。システムは、第1のネットワークに結合され、第1の情報ストリームを提供する、第1のフォーマットの送信元出力を有する。システムは、第2のネットワークに結合される、第2のフォーマットで受信すべき宛先入力も有する。宛先入力は、第2の情報ストリームを受信する。送信元出力と宛先入力との間に代理コード変換サーバ(「PTS」)が結合されている。代理コード変換サーバは、(様々な機能を有しうる)送信元端末の第1の機能を識別するように適合し、かつ(同様に様々な機能を有しうる)宛先端末の第2の機能を識別するように適合した、機能プロセスを有する。サーバは、1からN(Nは1より大きな整数)まで番号付けされた複数のコード変換モジュールを含むコード変換プロセスも有する。コード変換プロセスは、第1の機能および第2の機能に基づいて1つのコード変換プロセスを選択するように適合している。代理コード変換サーバは、ビット・レート制御プロセスを有する。ビット・レート制御プロセスは、第1のネットワークからネットワーク・ステータス情報(たとえば、ピング)を受信するように適合している。ビット・レート制御は、ネットワーク・ステータス情報に基づいて、情報ストリームのステータスを調整するようになっている(たとえば、停止、優先順位付けの許可、(下位ビット・レート・コードの選択による)ビット・レートの調整)。

【0013】

他の特定の態様では、本発明は、情報ストリームを処理する方法を提供する。この方法は、情報ストリーム用の複数の送信元機能から1つの送信元機能を識別する段階を含む。こ

の方法はまた、複数の宛先機能から1つの宛先機能を識別する。識別された送信元機能および識別された宛先機能に基づいて、ライブラリ内の複数のコード変換プロセスからコード変換プロセスを選択する段階が含まれる。この方法はまた、識別された送信元機能と識別された宛先機能が異なる場合に、選択されたコード変換プロセスを用いて情報ストリームを処理する。この方法はまた、識別された送信元機能と識別された宛先機能が一致する場合に、1つのコード変換プロセスとは無関係に情報ストリームを送信元から宛先に転送する。

#### 【0014】

本発明を用いると、従来の技術に勝る多数の利点をもたらされる。特定の態様では、本発明は、ビデオデータをH.263からMPEG-4データに、およびその逆にコード変換する（かつ 10  
他のビデオ・コーデック同士のコード変換を行う）か、またはオーディオ・データをG.723.1からGSM-AMRに、およびその逆にコード変換する（かつ他のオーディオ・コーデック同士のコード変換を行う）方法を提供する。好ましい態様において、コード変換は、コード変換されたデータを受信するエンド・ポイントが変換に気付かないようにシームレスに行われうる。本発明は、従来のソフトウェア技術およびデジタル信号プロセッサ（DSP）などのハードウェア技術を用いて実施することもできる。態様に基づいて、これらの利点または機能のうちの1つまたは複数を実現することができる。これらおよびその他の利点について本明細書全体にわたって説明し、以下に詳しく説明する。

#### 【0015】

本明細書に組み入れられ、かつ本明細書の一部を形成する添付の図面は、本発明の各態様 20  
を例示し、説明と共に、本発明の原則について説明する働きをする。

#### 【0016】

##### 特定の態様の説明

本発明によれば、電気通信分野におけるコード変換のための改良技術が提供される。特に、本発明は、ビデオ信号およびオーディオ信号をコード変換する方法および装置に関する。さらに、本発明は、少なくとも1つのプロセスが情報をコード変換するために選択される、複数のコード変換プロセスを有する代理コード変換サーバを用いて、情報（たとえば、ビデオ、オーディオ、データ）を第1のフォーマットから宛先フォーマットにコード変換する方法およびシステムを提供する。単なる例示のため、本発明は広域電気通信ネットワークに適用されるが、本発明を、インターネット、移動ネットワーク、ローカル・エリ 30  
ア・ネットワーク、PTSN、ISDN、SONET、DWDMなどのトランスポート・ネットワーク上で多数の異なる種類のマルチメディア・プロトコルを介して適用することもできることが認識される。

#### 【0017】

説明の一部は、様々な態様による、プログラミング命令を実行することによって行われる動作に関して提示される。当業者には理解されるように、このような動作は、たとえば電気構成要素によって記憶し、転送し、組み合わせ、かつその他の方法で処理することのできる、電気信号、磁気信号、または光学信号の形をとることが多い。説明の一部は、分散コンピューティング環境を用いて提示される。分散コンピューティング環境では、ファイル・サーバ、コンピュータ・サーバ、およびメモリ記憶装置をそれぞれの異なる場所に配 40  
置することができるが、これらの装置は、ネットワークを通じてローカル処理ユニットにアクセスすることができる。さらに、プログラム・モジュールは、種々のローカル・メモリ記憶装置およびリモート・メモリ記憶装置に物理的に配置することができる。各プログラム・モジュールは、ローカルでスタンドアロン方式で実行することも、リモートでクライアント・サーバ方式で実行することもできる。このような分散コンピューティング環境の例には、企業のローカル・エリア・ネットワーク、企業内コンピュータ・ネットワーク、およびグローバル・インターネットが含まれる。

#### 【0018】

さらに、本発明の各局面を説明するうえで読者を助けるために、以下の用語が与えられる。このような用語は、制限的なものではなく、当業者への説明のために与えられるに過ぎ 50



ない。当業者によって理解されている用語の意味に一致する、用語の他の意味も使用される。

用語	説明
ASIC	特定用途向け集積回路
CIF	共通中間フォーマット
ETSI	ヨーロッパ電気通信標準協会
G.723.1	ITU 勧告 G.723.1、5.3 kbit/s および 6.3 kbit/s で行われる マルチメディア通信用のデュアル・レート音声コーデ、1996 年
GOB	ブロック群
GSM	移動通信用グローバル・システム
GSM-AMR	ETSI アダプティブ・マルチレート音声コーデ。 GSM 06.90: 「デジタル・セルラー電気通信システム (フェーズ 2+); AMR 音声コード変換」、1998 年
GSM-AMR	GSM アダプティブ・マルチレート
H.320	ITU 勧告 H.320、狭帯域テレビ電話システムおよび端末機器、 1997 年

10

20

H.323	ITU 勧告 H.323、パケット・ベースのマルチメディア通信システム、1998年
H.324	ITU 勧告 H.323、低ビット・レート・マルチメディア、通信用端末、1998年
H261	ITU 勧告 H.263、 $p \times 64 \text{ kbit/s}$ でのオーディオビジュアル・サービス用ビデオ・コーデック、1993年
H263	ITU 勧告 H.263、低ビット・レート通信用ビデオ符号化、1998年
IETF	インターネット・エンジニアリング・タスク・フォース
ISO	国際標準化機構
ITU	国際電気通信連合
MB	マクロ・ブロック
MPEG	動画専門家団体、国際標準化機構の一部
MPEG2	MPEG オーディオビジュアル標準13818シリーズ
MPEG4	MPEG オーディオビジュアル14496 (1-5)
MVD	モーション・ベクトル・データ
P Frame or P Picture	予想される情報に基づくビデオ・フレーム
PTS	代理コード変換サーバ
QCIF	4分の1 CIF (CIF 参照)
RFC	コメントの要求
SDP	セッション記述プロトコル
SIP	セッション開始プロトコル
TCOEF or TCOEFF	変換係数
W3C	ワールド・ワイド・ウェブ・コンソーシアム
WAP	無線アクセス・プロトコル

## 【0019】

特定の態様では、「代理コード変換サーバ」（本明細書中ではPTS）という語は、現時点で既知である場合か、または既知でない、本明細書で説明する機能のいくつかまたはすべて、ならびにその他の機能を実行する様々なモジュールを有するコンピュータを指す。PTSは、ホスト・プロセッサ、1つまたは複数のネットワーク・インタフェース、および1つまたは複数のトランスコーダを含んでいる。トランスコーダは特に、プリント回路板、特定用途向け集積回路（ASIC）、およびフィールド・プログラマブル・ゲート・アレイ（FPGA）を含んでよい。PTSは、WAN、LAN、モバイル、PTSN、ISDN、SONETを含む、様々なネットワーク・アーキテクチャおよびプロトコルのコーデックおよびインタフェースを提供し、以下の機能のうちの1つまたは複数を実行することができる：

1. 機能の一致およびモードの選択；

2. メディア・ビットストリーム・コード変換；
3. メディア・ビットストリーム・レート制御；
4. 知的所有権管理および処理；
5. オーディオ・データ混合；および
6. 暗号化および／または復号。

【0020】

PTSの全体的な機能は、以下に例示する様々なプロトコル同士を変換することである。

【0021】

1. エンド・ポイントのメディア機能を搬送し、メディア・チャネルおよび会議を管理するのに用いられる、ITU H.32Xシリーズ、たとえばH.242やH.245を含む、マルチメディア・システム・プロトコルの変換。

【0022】

2. オーディオ・ストリームおよびビデオ・ストリームを含むメディア・ストリームの変換、たとえば、ビデオ・ストリームの、MPEG2、MPEG4、H.261、H.263を含む任意のビデオ・コーデック対の間の変換、またはオーディオ・ストリームの、G.723.1、G.729、GSM-AMR、EVRC、SMV、およびQCELPを含む任意のオーディオ・コーデック対の間の変換。

【0023】

3. IP管理プロトコルおよびIP権利マーキングの変換。たとえば、PTSは、MPEG4ストリームからIP権利情報を抽出し、メディアにおけるIP権利を維持するためのIP権利関連動作に従って情報を処理することができる。

【0024】

4. 必要に応じた信号の暗号化および／または復号。

【0025】

したがって、PTSは、様々なオーディオ変換機能およびビデオ変換機能を有している。機能の選択は、マルチメディア通信における全体的なサービス品質（QoS）に影響を与えるので、PTSが結合されるゲートウェイの接続帯域幅および負荷に基づいて決定されるべきである。PTSは、MPEGシリーズ、H.26Xビデオシリーズ、GSM-AMR、およびG.72Xオーディオ・コーデック・シリーズを含む様々なメディア・コンテンツに対するコード変換を行うことができる。さらに、PTSのコード変換機能は、そのプログラム可能性により、少なくとも部分的に容易にアップグレードすることができる。

【0026】

さらに、PTSはメディア・ビットストリーム・レート制御を行うことができる。レート制御が必要であるのは、2つのエンド・ポイント間の接続において、たとえば有線網から無線ネットワークへと、帯域幅が小さくなる可能性があるからである。PTSは、ネットワーク・アクセス・プロバイダから供給されたネットワーク・チャネル割当てから得た情報を用いるか、またはコマンド・プロトコルおよび制御プロトコル、たとえばH.242およびH.245を介してエンド・ポイントから発信することのできる帯域内帯域幅管理要求を用いて、レート制御を行うことができる。

【0027】

さらに、PTSは知的所有権（IP）管理および処理を行うことができる。たとえば、PTSは、知的所有権に関するデータ・セットを識別し、それを用いて管理および処理を容易にすることができる。一態様では、MPEG4は、コンテンツ、コンテンツの種類、およびIP権利保有者に関する情報を保持する任意の知的所有権識別（IPI）データ・セットで、符号化されたメディア・オブジェクトを補足する。データ・セットが存在する場合、それはメディア・オブジェクトに関連するストリーミング・データを記述する、基本ストリーム記述子の一部を形成する。各メディア・オブジェクトに関連するデータ・セットの数は様々であってよく、いくつかの異なるメディア・オブジェクトが、同じデータ・セットを共用することができる。データ・セットを与えることによって、証跡、監視、課金、およびコピー保護に関する機構を実施することができる。

【0028】

マルチメディア通信用途は、IP権利保護およびセキュリティに対する広範囲の要件を有する。用途によっては、情報が本質的な価値を有さない場合でも、プライバシーを維持するためにユーザが交換する情報を保護する必要がある。他の用途では、情報の作成者および／または配給者に対する重要な情報に関する高度の管理および保護を必要とする。さらに、IP権利管理および処理のフレームワークは一般に、特定の用途に必要とされ、かつ特定のビットストリーム内に格納される、様々な形式のIP権利データにアクセスできるような融通性を有さなければならない。

#### 【0029】

以下の説明では、ゲートウェイにスタンドアロン装置として接続されたPTSを示すが、PTSの他の態様は、ネットワークに接続された交換機、サーバ、ルータ、または任意の装置に接続されたPTSを使用することができる。さらに、PTSは、ネットワークに接続されたゲートウェイ、交換機、ルータ、または任意の装置に、そのネットワーク装置の一部を形成するように組み込むことができる。他の態様の詳細を以下に提供する。

#### 【0030】

図2は、PTSが移動エンド・ポイントとLAN電話との間の信号をコード変換する態様を示している。この図は一例に過ぎず、特許請求の範囲を過度に制限するものではない。当業者には、他の多数の変形態様、修正態様、および代替態様が認識される。移動エンド・ポイント210は、LAN電話220の間でオーディオ情報およびビデオ情報を送受信する。送信された情報は、移動エンド・ポイント210から無線通信チャネル、たとえばエア・リンクを介して基地局230に至り、その後マスタ・サービス・コントローラ240、ニア・エンド・ゲートウェイ250に至り、インターネット260を通じてファアエンド・ゲートウェイ270、LAN交換機280に至り、最終的にLAN電話220に至る。

#### 【0031】

移動エンド・ポイント210のエア・リンクおよび移動性の制限によって、移動エンド・ポイント210と基地局230との間の帯域幅は、インターネット・ルータとLAN電話220との間の帯域幅よりもずっと小さい数十Kbpsまたは数百Kbpsに達しうる。後者の帯域幅は毎秒数十メガビット (Mbps) から数百メガビットに達する。図2で、移動エンド・ポイント210のメディア符号化・復号機能は、オーディオ信号の場合はGSM-AMR、およびビデオ信号の場合はMPEG-4である。これに対して、LAN電話230のメディア符号化・復号機能は、オーディオの場合はG.723.1であり、ビデオの場合はH.263である。機能および帯域幅の差により、移動エンド・ポイント210とLAN電話230との間の経路において、PTSによって重要なコード変換が行われ、ほぼリアルタイムの通信が容易になる。さらに、PTSによるトランスコーディングは、PTSがスロットリング (throttling) 機能を実行するため、移動ネットワーク290の過負荷を防止する。

#### 【0032】

図2では、オーディオ信号をG.723.1とGSM-AMRとの間でコード変換する必要がある、それに対して、ビデオ信号をMPEG4ビデオとH.263との間でコード変換する必要がある。このようなコード変換は、移動ネットワーク290でPTS204によって行うことも、陸上通信線ネットワーク292でPTS208によって行うこともできる。したがって、コード変換を行ううえで2つのPTS204および208は必要とされない。しかし、各ネットワークで1つのPTSを実施すると、各ゲートウェイにおける帯域幅要件を低くすることができる。さらに、図2に示されている態様がPTSによるコード変換を示しているにもかかわらず、PTSはコード変換せずに信号を送信することも可能である。

#### 【0033】

図3は、本発明の他の態様のブロック図である。この図は、一例に過ぎず、特許請求の範囲を過度に制限するものではない。当業者には、他の多数の変形態様、修正態様、および代替態様が認識される。モバイル・ハンドセット310は、ビデオ・コンテンツ・サーバ320との間で情報を送受信する。ビデオ・サーバ320は、オーディオ信号およびビデオ信号を含む映画をストリーミングする。オーディオ信号は、MPEG2オーディオ・レベル3 (MP3) を用いて符号化され、ビデオ信号は、MPEG2ビデオを用いて符号化される。ビデオ・サー

バ320とモバイル・ハンドセット310とで帯域幅および機能が一致しないため、移動ネットワーク390内のPST304は、オーディオをGSM-AMRにコード変換し、かつビデオをMPEG4ビデオにコード変換する。しかし、ビデオ・サーバ・ネットワーク392内のPST308がコード変換を実行する場合、モバイル・ハンドセット・ネットワーク390内のゲートウェイに対する帯域幅要件を低くすることができる。

#### 【0034】

他の態様では、通信の2つのエンド・ポイントの機能をWAP (Wireless Access Protocol) ユーザ・エージェント・プロファイル情報 (機能および嗜好情報)、ウェブ・コンソーシアムCC/PP (リソース記述フレームワーク、すなわちRDFを用いる)、IETF標準 (RFC2506、RFC2533、およびRFC2703)、またはITUのH.245標準もしくはH.242標準、あるいはそれらの組合せを用いて指定することができる。他の態様において、ゲートウェイは、PTSの支援の下で、エンド・ポイントの機能、およびメディアの送信に利用可能か、または割り当てられた帯域幅を検出することができる。次いで、PTSは、データ符号化モードを選択し、エンド・ポイントの要件を最もよく満たすようにメディアをコード変換することができる。したがって、PTSは、エンド・ポイント側の代理として働き、一方の側に適切なある形式のビットストリームを、他方の側に適切な他の形式に変換することができる。

#### 【0035】

図4は、本発明の態様によるPTSのブロック図を示している。この図は、一例に過ぎず、特許請求の範囲を過度に制限するものではない。当業者には、他の多数の変形態様、修正態様、および代替態様が認識される。論理ポートは抽象通信ポートとして定義されている。1つの物理的ネットワーク・インタフェース上にいくつかの論理ポートが存在してよいが、PTSは複数の物理的ネットワーク・インタフェースを有してよい。ネットワーク・ゲートウェイ・コントローラ・インタフェース (NGCI) 420は、PTS410とゲートウェイまたはネットワーク・ホストとの間でメッセージを送受信するための、少なくとも1つのNGCI論理ポートを含んでよい。ゲートウェイはエンド・ポイント同士の間で呼シグナリング (call signaling) を実行することができる。ゲートウェイは、メディア・ゲートウェイ・コントローラであっても、エンド・ポイント同士の間で呼またはトランスポート機能確立する、任意の他のゲートウェイ装置であってもよい。呼シグナリングは、コード変換すべきメディア・ストリームの送信元、および宛先のネットワーク・アドレスを交換することによって、エンド・ポイント・エンティティ同士の間で初期リンクを確立するプロセスである。リアルタイム・インタフェース・プロトコルの場合、呼シグナリングは、メディア・ストリームの送信元および宛先の、インターネット・プロトコル・アドレスおよびインターネット・プロトコル・ポート番号を確立する。呼シグナリングは、SIPまたはH.323によって必要とされるような、より精密なプロセスを含んでもよい。

#### 【0036】

メディア・ネットワーク・インタフェース (MNI) 430は、メディア・ビットストリームの送受信の論理ポートを提供する。これらの論理ポートを通じて、PTS410はビットストリーム、たとえば、オーディオ信号、ビデオ信号、コマンドおよび制御データ、ならびにテキストであってもバイナリであってもよいその他のデータを受信する。PTS410によってビットストリームを受信するためのポート・アドレス、およびコード変換されたストリームが送信される宛先アドレスは、NGCI420を通じて呼シグナリング・ゲートウェイとPTS410との間のメッセージング・プロトコルによって指定される。MNI430の基本的な物理的インタフェースは、ギガビット・ネットワーク・インタフェース・カード、E1/T1/OC3などの時間ドメイン多重化 (TDM) 回線交換接続であってよい。MNI430物理的ネットワーク・インタフェースは、NGCI420の物理的リンクを共用してもしなくてもよい。

#### 【0037】

モニタ・セットアップ・インタフェース (MSI) 440は、PTS410の初期環境設定、監視、および再環境設定に使用される。MSIデータが搬送される物理的リンクは、他のネットワーク接続と共用するか、または専用シリアル接続ポートを通じて搬送することができる。ゲートウェイ、たとえばメディア・ゲートウェイ・コントローラとの接続、およびコンテン

ト・サーバ、たとえばビデオ・サーバとの接続を含む、PTSに関する多数の接続性が存在する。

#### 【0038】

図5は、本発明の態様によるPTSの接続性を示す図である。この図は、一例に過ぎず、特許請求の範囲を過度に制限するものではない。当業者には、他の多数の変形態様、修正態様、および代替態様が認識される。MNI530は、PTS510をルータまたは交換機540を介してインターネット550に接続し、NGCIは、PTS510をNGCI520を介してコンテンツ・サーバ560に接続する（またはゲートウェイ560は、2つのエンド・ポイント、たとえば携帯電話とIP電話を仲介するゲートウェイであってもよい）。コンテンツ・サーバが基本的な端末エミュレーション・サポート、たとえば、Windows下でのハイパーターム（hyper-term）を提供できる場合、MNI530をコンテンツ・サーバ560に接続してもよい。一態様では、ゲートウェイまたはコンテンツ・サーバはネットワーク・ホストであってもよい。

#### 【0039】

図6は、本発明の態様に従う動作サイクルによる、様々なPTS機能を示している。この図は、一例に過ぎず、特許請求の範囲を過度に制限するものではない。当業者には、他の多数の変形態様、修正態様、および代替態様が認識される。ネットワーク、たとえばネットワーク・ホストやルータにPTSが接続され、電力が投入され、環境設定されると、PTSは、ネットワーク・ホストからコード変換セッションの開始を受け入れる準備を完了する。

#### 【0040】

段階606で、PTSはゲートウェイまたはコンテンツ・サーバからメッセージを受信する。段階608で、PTSは、受信されたメッセージがセッション・メッセージであるかどうかを判定する。メッセージがセッション・メッセージである場合、段階632で、PTSは、メッセージがセッション開始メッセージであるかどうかを判定する。MGCPプロトコルまたはH.248/MEGACOプロトコルのような、NGCI上で動作するメッセージング・プロトコルを通じて、ネットワーク・ホストによってセッションを開始することができる。1つまたは複数のセッションが任意の時間に到着することができる。セッションが開始された後、このセッションに関するPTSの基本動作はさらなるメッセージによって制御される。

#### 【0041】

受信されたメッセージがセッション開始メッセージである場合、段階650でPTSは新しいセッションを開始する。受信されたメッセージがセッション開始メッセージではない場合、段階634で、PTSは、受信されたメッセージがセッション維持メッセージであるかどうかを判定する。メッセージがセッション維持メッセージである場合、PTSは段階652でセッション維持プロセスを開始する。セッション維持プロセスは、セッション用の接続の維持または終了を含んでよい。受信されたメッセージがセッション維持メッセージである場合、段階636で、PTSは、受信されたメッセージがコード変換メッセージであるかどうかを判定する。メッセージがコード変換メッセージである場合、段階654で、PTSはコード変換メッセージを処理する。メッセージがコード変換メッセージではない場合、段階638で、PTSは、受信されたメッセージがレート制御メッセージであるかどうかを判定する。受信されたメッセージがレート制御メッセージである場合、段階656で、レート制御メッセージが処理される。レート制御プロセスは、PTSに、送信帯域幅を効率的に利用し、かつ送信されたパケットが脱落するのを防止するために、送信速度を動的に調整するよう指示する。

#### 【0042】

PTSがビットストリームをコード変換する際、PTSによって生成される最適なデータ・レートは、PTSとビットストリームの宛先との間のネットワークのメディア・プロトコルに依存する。宛先によって受信されるデータ・レートは、ネットワーク輻輳、ルータと宛先との間のリンクの種類、たとえば有線接続または無線接続、リンクに関連するプロトコルおよびデータ多重化、ならびにリンクの品質に応じて異なる。宛先は、処理できるデータ・レートよりも高いデータ・レートを受信し、バッファ・オーバーフローを起こすことがある。一方、宛先は、受信すると予期したデータ・レートよりも低いデータ・レートを受信し、バッファ・アンダーフローを起こすことがある。したがって、PTSは、バッファ・オー

バッフローとバッファ・アンダーフローの両方を避けるようにそのデータ・レートを調整する必要がある。

#### 【0043】

PTSは、少なくとも以下の方法によってそのデータ・レートを調整する。PTSは、ネットワーク・ホストもしくはネットワーク・アクセス・プロバイダ、または内部PTS機構から得たネットワーク輻輳情報、帯域幅情報、品質情報を用いて往復時間を算出する。PTSとエンド・ポイント、たとえば、ビットストリームの送信元または宛先との間の往復時間は、このエンド・ポイントに「ピング」を送信することによって測定することができる。「ピング」パケットがエンド・ポイントに到着するまでの総時間と、応答パケットがエンド・ポイントからPTSに到着するまでの総時間が往復時間である。ネットワークが輻輳すればするほど、往復時間は長くなる。したがって、往復時間を用いて現在のPTSビット・レートでのネットワークの輻輳レベルを評価することができる。

#### 【0044】

または、PTSは帯域内情報を用いてネットワーク輻輳を評価することができる。たとえば、PTSは、H.324およびH.323で使用されるH.245などのプロトコルの下で、PTSのビット・レート・スループットを減らすかまたは増やす命令を受信することができる。

#### 【0045】

瞬間的なネットワーク条件が与えられた場合に、帯域内法または往復時間法を用いて適切なサービス品質を維持することができる。PTSは、以下の方法を用いて適切なビット・レートを生成する適切な符号化モードを、輻輳情報およびビット・レート情報を用いて判定する。

#### 【0046】

一態様では、PTSは、リアルタイム動作におけるサービス目標を満足させるために、コード変換パラメータを変更することができる。たとえば、MPEG4ビデオでは、低ビット・レート・スループットを生成するために量子化パラメータを変更することができる。しかし、この変更はビデオ品質の低下を導きうる。したがって、ビデオ品質が重要である場合、H.263およびMPEG4によって提供されるような高度な符号化技術を用いて、品質を低下させずにビット・レート・スループットを減らすことができる。しかし、これらの方法は、より高い計算要件が課される可能性がある。したがって、PTSレート制御方式は、信号品質とビット・レートと計算との所望のバランスをとる必要がある。

#### 【0047】

たいていのオーディオ・コーデックは、微細な可変レートを与えることができないが、オーディオにおけるレート調節は同様に行われ、その代わりに、PTSが選択することのできるいくつかのビット・レートを提供する。たとえば、G.723.1オーディオ・コーデックは2つのビット・レート、すなわち低いレートと高いレートを提供する。同様に、GSM-AMRコーデックは、4.75Kbpsから12.2Kbpsの範囲の8つのビット・レートをサポートする。PTSは、このエンド・ポイントまでのネットワーク経路が輻輳しているか、またはこのエンド・ポイントを含むリンクに割り当てられた帯域幅が低帯域幅である場合、より低いレートを使用することができる。

#### 【0048】

他の態様では、ルータなどのネットワーク機器がデータ供給の優先順位付けをサポートする場合、PTSは、PTSによって取り扱われているデータにより高い優先順位を与えるようそのネットワーク機器に命令することができる。たとえば、インターネット・プロトコルのバージョン6は、パケットの優先順位付けをサポートする。さらに、IETFは、各エンド・ポイントが帯域幅を確保するのを可能にするリソース確保に関する標準を開発している。PTSが配備されているネットワークによって、パケット優先順位付けおよびリソース確保がサポートされている場合、PTSはこれらを利用することができる。たとえば、インターネット・プロトコルは、パケットに優先順位を割り当てる機能を提供し、PTSはこの機能を用いて、必要に応じてパケットを優先順位付けすることができる。さらに、より高い優先順位の接続が、実施可能になった直後に処理されるように、PTSは、PTS自体の接続の内

部優先付けをサポートすることができる。

【0049】

再び図6を参照すると、段階640で、PTSは受信されたメッセージが機能メッセージであるかどうかを判定する。機能メッセージとは、各エンド・ポイントの機能を含むメッセージである。メッセージが機能メッセージである場合、段階658で、PTSはこのメッセージを処理する。PTSは、エンド・ポイントの機能を定義するメッセージを処理することができる。たとえば、エンド・ポイントがビデオサーバおよび移動端末である場合、ビデオサーバおよび移動端末の機能がPTSに搬送される。したがって、PTSは、この2つのエンド・ポイント間の通信の最良のモードを判定する。最良のモードの選択において、移動ユーザが見たい特定のビデオコンテンツに関連するプロトコルを考慮する。様々な種類のコンテンツを異なるプロトコル、たとえばMPEG2およびMPEG4によって符号化することができる。さらに、移動端末の機能を多数の方法でネットワーク・ホストを介してPTSに送信することができる。たとえば、PTSは、移動端末に記憶されている情報から、ユーザのサービス・プロバイダのネットワーク・データベースに記憶されているユーザ加入情報から、または呼シグナリング・フェーズ中に移動端末とネットワーク・アクセス・ゲートウェイとの間で交換されるビットストリーム内の帯域内情報から、機能を得ることができる。機能のフォーマットには特に、ITU、IETF、およびWAPが含まれる。

【0050】

機能メッセージは、一方のエンド・ポイント、たとえばビデオサーバから、他方のエンド・ポイントに送信される特定のメディアのための最良のコード変換モードを判定するために、ネットワーク・ホストからPTSに送信されうる。機能モード選択プロセスにおいて、PTSは、送信元からデータを受信するためのあるビットストリーム・プロトコル・モード、および受信されたメディアをPTSが変換する際の変換先となる、他のビットストリーム・プロトコル・モードを選択することができる。

【0051】

他の態様では、ビットストリーム・トランスポート・チャネルを開くために、各エンド・ポイントごとの選択されたモードを、それぞれのエンド・ポイントに通知することができる。H.323またはH.324では、ネットワークはH.245論理チャネル動作または高速接続手順を用いてこのようなチャネルを開く。H.245論理チャネル動作を用いて、エンド・ポイントは、信号を送信するために他方のエンド・ポイントに「論理チャネル開放」要求を送信することができる。H.323において、エンド・ポイントは、たとえば、H.225.0によって勧告されたITU Q.931標準の下で、呼の初期設定時に交換される呼シグナリング情報にカプセル化される「高速開始」メッセージを用いることによって、信号の送信準備を完了するためのメディア・チャネルについての情報をカプセル化することができる。したがって、各エンド・ポイント用の選択されたプロトコルが与えられると、ネットワーク・ホストまたはPTSは、メディア・ビットストリーム用のトランスポート・チャネルを確立することができる。開放プロセスは、ネットワークによって仲介されるエンド・ポイント間の全体的な接続の、システム・レベル・プロトコルに依存する。

【0052】

PTSがコード変換を行うには、選択されたメディア送信モードを送信元アドレスおよび宛先アドレスに関連付ける必要があり、かつそのような関連付けの情報をメッセージを介してPTSに伝達する必要がある。この関連付けは、H.245標準で行われるように、一方のエンド・ポイントによって論理メディア・チャネルを開くことによって得られる。PTSがメディア送信モードの種類を選択したときの、論理メディア・チャネルのこのような暗黙の開放は、ゲートウェイもしくはコンテンツ・サーバによって明示的に要求するか、またはいくつかの標準要件の下で事前にプログラムすることができる。送信元アドレスまたは宛先アドレスを指定する手段にかかわらず、選択されたメディア送信モードと、送信元アドレスまたは宛先アドレスの関連付けによって、入力ビットストリームを得るべき場所、およびコード変換されたビットストリームを送信すべき場所が、PTSに通知される。

【0053】

10

20

30

40

50



特定の態様では、PTSは、送信元アドレスからビットストリームを読み取り、ビットストリームをその最初のフォーマットから目標フォーマットにコード変換し、変換されたビットストリームを宛先アドレスに送信する。ビットストリーム・データの送受信は、ネットワーク・ハードウェア固有のソフトウェアを用いて、ネットワーク読取り／書き込み機能によって行われる。

#### 【0054】

再び図6を参照すると、段階642で、PTSは、受信されたメッセージがネットワーク・アドレス指定メッセージであるかどうかを判定する。ネットワーク・アドレス指定メッセージは、ビットストリームの送信元および／または宛先のネットワーク・アドレスに関する情報を含む。受信されたメッセージがネットワーク・アドレス指定メッセージである場合、段階660で、PTSはこのネットワーク・アドレス指定メッセージを処理する。段階644で、PTSは、受信されたメッセージがメディア混合メッセージであるかどうかを判定する。メディア混合メッセージとは、2つまたはそれ以上のオーディオ・ストリームに関連する信号を混合し、混合されたビットストリームをネットワーク宛先アドレスに再送信するよう、PTSに要求するメッセージである。受信されたメッセージがメディア混合メッセージである場合、段階662で、PTSはこのメディア混合メッセージを処理する。段階646で、PTSは、受信されたメッセージがIP権利メッセージであるかどうかを判定し、そうである場合、段階664で、PTSは、受信されたメッセージに含まれているIP権利に関する情報を、このメッセージ中の命令に従って管理する。いくつかのメディア通信・表現プロトコルは、IP権利管理および処理とのインタフェースと、ビットストリームに含まれているIP権利に関する情報へのアクセスをサポートする。たとえば、PTSは、IP権利管理および処理に関する、MPEG-4インタフェース仕様をサポートすることができる。IP権利に関する情報は、MPEG4ビットストリームから抽出または逆多重化され、メッセージング・システムを介してネットワーク・ホストが利用できるようにされる。さらに、ネットワーク・ホスト上にインストールされているか、またはPTSに挿入されているIP権利固有のアプリケーションは、記録維持、コンテンツの再署名、およびブロックングを含む様々な目的で、このような情報にアクセスし、かつ処理する。

#### 【0055】

図6の段階610で、PTSは、受信されたメッセージが機能・モード・メッセージであるかどうかを判定し、そうである場合、段階622で、PTSはこのメッセージを処理する。機能・モード・メッセージの処理には、コード変換、混合、ならびにセッションに関連するその他のオプションに関係する機能およびオプションの起動が含まれる。段階612で、PTSは、受信されたメッセージがリソース・メッセージであるかどうかを判定し、そうである場合、段階624で、PTSはこのメッセージを処理する。リソース・メッセージの処理には、PTSが管理するハードウェア処理リソース、メモリ・リソース、およびその他のコンピューティング・リソースまたはネットワーク化リソースに関するPTSリソースの処理が含まれる。PTSが、受信中であり後で送信されるメディアを暗号化する必要があるという命令を受信した場合、PTSは送信されるデータをコード変換した後、暗号化する。同様にPTSは、受信した命令に応じてデータを復号しうる。

#### 【0056】

上記では、PTSの機能が分離した段階として示されているが、当業者には、これらの分離した段階のうちの1つまたは複数を組み合わせるか、またはさらに細分して分離した段階の機能を実行できることが理解されることが考えられる。態様に応じて、前述の機能を分離することも、場合によって組み合わせることもできる。機能は、任意の組合せを含むソフトウェアおよび／またはハードウェアで実施することができる。態様に応じて、他の多数の修正態様、変形態様、および代替態様が存在してよい。

#### 【0057】

図7、8、9、および10は、主要なシステム・メッセージの態様を示す、簡略化された流れ図である。これらの図は、一例に過ぎず、特許請求の範囲を過度に制限するものではない。当業者には、他の多数の変形態様、修正態様、および代替態様が認識される。主要なシ

システム・メッセージは、PTSがメディア・ゲートウェイ・コントローラなどのネットワーク機器からの命令に応答するのを可能にする。メッセージには、セッションを開始する命令、セッションを終了する命令、セッション・オプションを設定する命令、セッション・オプションを得る命令、セッション・マネージャにメッセージを送信する命令、セッション・マネージャからメッセージを得る命令、PTSモードを設定する命令、PTS機能を設定する命令、PTS機能を得る命令、リソース・ステータスを得る命令、リソース・ステータスを設定する命令、ファームウェア手順を更新する命令、PTSシステム・ステータスを得る命令、PTSをリセットする命令、PTSをシャットダウンする命令、およびデバック／追跡モードを起動する命令が含まれる。

#### 【0058】

10

図11~16は、PTSセッション維持メッセージおよびコード変換メッセージの態様を示す、簡略化された流れ図である。これらの図は、一例に過ぎず、特許請求の範囲を過度に制限するものではない。当業者には、他の多数の変形態様、修正態様、および代替態様が認識される。セッション維持メッセージおよびコード変換メッセージは、PTSが、セッションを終了し、コード変換チャネルを開き／閉じ、コード変換オプションを設定し／得て、エンド・ポイント機能を更新し、エンド・ポイント機能を一致させ、エンド・ポイント機能を選択し、レート制御を起動し、レート制御モードを得て、メディア宛先アドレスを追加／削除し、メディア送信元アドレスを追加／削除し、IPRモードを設定し／得て、IPRオプションを設定し、メディア・チャネルの混合を起動し、混合を無効にし、かつチャネル混合モードを設定するのを可能にする。

20

#### 【0059】

図17は、本発明の態様に従ってPTSで使用するこことできる、ソフトウェア・モジュールを示すブロック図である。この図は、一例に過ぎず、特許請求の範囲を過度に制限するものではない。当業者には、他の多数の変形態様、修正態様、および代替態様が認識される。PTSソフトウェアは以下の主要モジュールを含んでいる：

1. セッション管理モジュール1710；
2. PTS管理モジュール1720；
3. ネットワーク・ホスト・インタフェース・モジュール1730；
4. メディア・チャネル処理モジュール1740；
5. 呼シグナリング・インタフェース・モジュール1750；
6. ネットワーク・インタフェース・モジュール1760；
7. コード変換モジュール1770；
8. レート制御モジュール1780；
9. 知的所有権管理モジュール1790；および
10. 機能処理モジュール1792。

30

#### 【0060】

セッション管理モジュール1710は、ゲートウェイの主要サービスを実行し、したがって、これは主要PTSソフトウェア・プログラムである。たとえば、セッション管理モジュール1710は、コード変換セッションを開始および終了し、セッション・メッセージを処理および送出し、かつセッション・リソースを管理する。PTS管理モジュール1720は、オペレータがPTSのステータスを検査するか、またはPTSの領域ソースを管理する必要がある場合、必要な基本的で全体的な管理機能を実行する。たとえば、管理モジュール1720は、構成要素固有の試験手順によって主要ハードウェア構成要素を試験し、PTSをリセットし、コード変換リソースを動的に追跡し割り当てる。ネットワーク・ホスト・インタフェース・モジュール1730は、PTSとネットワーク・ホスト、たとえば、メディア・ゲートウェイ・コントローラまたはコンテンツ・サーバとの間の通信メッセージング・インタフェースを処理する。たとえば、インタフェース・モジュール1730は、ネットワーク・ホストの種類に応じて、ネットワーク・ホストとPTSとの間のメッセージングを実施し、コード変換機能をPTSが取り込むか、またはPTSに関して定義するための、方法定義コマンドを実施する。さらに、インタフェース・モジュール1730は、機能交換が明示的に行われな

40

50

Sがメディア・コンテンツ・タイプをビットストリームから検索するための方法を実施することができる。メディア・チャネル処理モジュール1740は、チャネル・ネットワーク送信元および宛先の開閉、追加、削除などのメディア・チャネル機能を実行する。呼シグナリング・インタフェース・モジュール1750は、PTSを通じてエンド・ポイント同士の間の初期呼設定を確立する機能を実行し、ここで呼設定の手順はSIPおよびQ.931などの標準に依存する。ネットワーク・インタフェース・モジュール1760は基本入力および/または出力通信インタフェースを実現する。基本入力および/または出力は、より複雑なメッセージングが行われる最低レベルの通信である。

#### 【0061】

コード変換モジュール1770は、MPEGシリーズ、H.26Xビデオシリーズ、GSM-AMRおよびG.72 10  
Xオーディオ・コーデック・シリーズ間のコード変換を含む、実際のコード変換機能を実行する。PTSコード変換の他の例には、少なくともMPEG2オーディオからMPEG4オーディオ、G.723.1からGSM-AMR、MPEG2ビデオからMPEG4ビデオ、H.263からMPEG4ビデオへのコード変換を含んでよい。レート制御モジュール1780は、スロットリング機能を実行し、エンド・ポイント同士が通信するネットワーク・セグメントが過負荷状態になるのを防止する。知的所有権モジュール1790はIP権利を保護する。たとえば、管理モジュール1790は、ビットストリーム中のIP権利に関するデータを使用し、受信されコード変換されたビットストリームに関連するIP権利の監査、監視、課金、および保護を行う実施機構を助ける。機能処理モジュール1792は、一方のエンド・ポイントから別のエンド・ポイントに送信される特定のメディアに最もよく一致するモードを、受信されたメッセージ中のデータを利用し 20  
て見つける。

#### 【0062】

PTSのアーキテクチャは、サーバの性能、コスト、および製品化までの時間を決定する。性能は、PTSが同時に処理することのできる同時ゲートウェイ・チャネルまたは呼の数とみなすことができる。さらに、一定数のチャネルについて、アーキテクチャのコストおよび性能は以下の因子に依存する：

- 1.バス・アーキテクチャ；
- 2.様々なビデオおよびオーディオ・コード変換用のコード変換アーキテクチャおよびハードウェア；
- 3.MGCおよびその他のゲートウェイ構成要素に接続するためのネットワーク・オフ・ロー 30  
ディング；
- 4.オペレーティング・システム。

#### 【0063】

図18は、フローチャートで用いられる記号の説明を示している。この図は一例に過ぎず、特許請求の範囲を過度に制限するものではない。当業者には、他の多数の変形態様、修正態様、および代替態様が認識される。

#### 【0064】

図19は、本発明の一態様に従ってPTSで使用されうる、ビデオ・ビットストリームコード変換の高レベル手順を例示する、完結化された流れ図を示す。この図は一例に過ぎず、特許請求の範囲を過度に制限するものではない。当業者には、他の多数の変形態様、修正態 40  
様、および代替態様が認識される。この手順では、ビットのチャンクが読み取られ、エンド・シーケンス・マーカが検出された場合、手順は終了する。エンド・シーケンス・マーカが検出されない場合、手順では次の符号語が読み取られ、符号語が出力プロトコル符号語にコード変換され、履歴レコードが更新され、コード変換された各ビットは、受信側エンド・ポイントにある入力バッファがオーバーフローしないようにレート制御スキームに従ってフラッシュされる (flushed)、出力バッファへ出力される。

#### 【0065】

図20は、本発明の一態様によるPTSハードウェア・アーキテクチャを示す。この図は一例に過ぎず、特許請求の範囲を過度に制限するものではない。当業者には、他の多数の変形態様、修正態様、および代替態様が認識される。このアーキテクチャは、1つまたは複数 50

のプロセッサ、すなわち、DSPプロセッサ2020のバンク、1つまたは複数のネットワーク・インタフェース2030、1つまたは複数のプロセッサ2040、およびメモリ・バンク2050を含むバス・カードであるインテリジェント・コード変換ノード2010を備えている。このアーキテクチャはいくつかの利点を有している。まず、インタフェース2030はインテリジェント・コード変換ノード・バス・カード2010に埋め込まれている。したがって、呼処理およびコード変換は、バス・カード2010上でローカルに行われる。第2に、1つまたは複数のネットワーク・インタフェースが可能である。第3に、このアーキテクチャはその処理方法がコンパクトであるため多数の同時呼をサポートすることができる。

【0066】

図20の態様に加えて、PTSの多数のアーキテクチャが可能であり、以下にそれらのうちのいくつかを列挙する：

- 1.バス・カードを有するスタンドアロン・シャシ；
- 2.後述のPC様実施態様；
- 3.ASICを含む、既存の処理ハードウェアのファームウェア；
- 4.既存のハードウェア上で実行されるソフトウェア；
- 5.ASIC、DSP、または他の種類のプロセッサによってハードウェアを加速する、既存のハードウェア上で実行されるソフトウェア；および
- 6.ASICチップセット。

【0067】

図21は、本発明によるコンピュータ・システムの一態様を示している。この図は一例に過ぎず、特許請求の範囲を過度に制限するものではない。当業者には、他の多数の変形態様、修正態様、および代替態様が認識される。本発明は、パーソナル・コンピュータ（PC）アーキテクチャで実施することができる。他のコンピュータ・システム・アーキテクチャ、または他のプログラム可能な装置もしくは電子ベースの装置を使用することもできる。

【0068】

図21で、コンピュータ・システム2100は、情報を送信するバス2101、バス2101に結合され、情報を処理するプロセッサ2102、バス2101に結合され、プロセッサ2102のための情報および命令を記憶するランダム・アクセス・メモリ2103、バス2101に結合され、プロセッサ2102およびPTSアプリケーションのための静的な情報および命令を記憶する読取り専用メモリ2104、バス2101に結合され、ユーザのための情報を表示する表示装置2105、バス2101に結合され、情報およびコマンド選択をプロセッサ2102に送信する入力装置2106、バス2101に結合され、情報および命令を記憶する、磁気ディスクおよび関連するディスク・ドライブなどの大容量記憶装置2107とを含んでいる。データ記憶媒体2108は、デジタル情報、たとえば、PTSソフトウェア・モジュールを含み、かつバス2101に結合され、データ記憶媒体2108に記憶されているデジタル情報への、バス2101を通じたプロセッサ2102のアクセスを提供するために、大容量記憶装置2107と協働するように構成されている。ハードウェア・コード変換加速モジュール2109は、プリント回路板、デジタル信号プロセッサ、ASIC、およびFPGAを含んでいる。モジュール2109は、バス2101に通信可能に結合されており、図20に示されているように、インテリジェント・コード変換ノード2010と同様の態様を有することができる。

【0069】

プロセッサ2102は、様々な汎用プロセッサまたはマイクロプロセッサ、たとえば、Intel Corporationによって製造されているPentium（商標）プロセッサ、および94039-7311カリフォルニア州マウンテンビューN.ショアライン通り2011のMIPS Technologies, Inc.によって製造されているMIPSプロセッサであってよい。デジタル信号プロセッサ（DSP）のような他の様々なプロセッサをコンピュータ・システム2100で使用することもできる。表示装置2105は、液晶装置、陰極管（CRT）、またはその他の適切な表示装置であってよい。大容量記憶装置2107は、ハード・ディスク、フロッピー・ディスク、CD-ROM、磁気テープ、またはその他の磁気データ記憶媒体もしくは光学データ記憶媒体に記憶されている情報を読み書きする、従来のハード・ディスク・ドライブ、フロッピー・ディスク・ドライブ

、CD-ROMドライブ、またはその他の磁気データ記憶装置もしくは光学データ記憶装置であってよい。データ記憶媒体2108は、ハード・ディスク、フロッピー・ディスク、CD-ROM、磁気テープ、またはその他の磁気データ記憶媒体もしくは光学データ記憶媒体であってよい。

#### 【0070】

一般に、プロセッサ2102は、読取り専用メモリ2104から処理命令およびデータを取り込むことができる。プロセッサ2102は、大容量記憶装置2107を用いてデータ記憶媒体2108から処理命令およびデータを取り込むこともでき、SDRAMであってよいランダム・アクセス・メモリ2103に情報をダウンロードする。プロセッサ2102は次いで、ランダム・アクセス・メモリ2103または読取り専用メモリ2104からの命令ストリームを実行する。入力装置2106 10  
に入力されたコマンド選択および情報は、プロセッサ2102によって実行される命令の流れを指示することができる。入力装置2106は特に、従来のマウスやトラックボール装置などのポインティング装置であってよい。実行結果は表示装置2105上に表示することができる。コンピュータ・システム2100は、コンピュータ・システム2100をネットワークに接続するネットワーク装置2110も含んでいる。ネットワーク装置2110は、イーサネット装置、電話ジャック、衛星リンク、またはその他の装置であってよい。

#### 【0071】

本発明の態様は、機械アクセス可能媒体上に記憶され、コンピュータ・アクセス可能媒体またはプロセッサ・アクセス可能媒体とも呼ばれるソフトウェア製品として表すことができる。機械アクセス可能媒体は、ディスケット、CD-ROM、揮発性であってても非揮発性であ 20  
ってもよいメモリ装置、ASIC、ASICのファームウェア、システム・オン・チップ、またはその他の記憶機構を含む、任意の種類の磁気記憶媒体、光学記憶媒体、または電気記憶媒体であってよい。機械アクセス可能媒体は、命令、符号シーケンス、または構成情報の様々なセットを含んでよい。本発明を実施するのに必要な他のデータを、機械アクセス可能媒体上に記憶することもできる。単なる一例として、コード変換技術は、共通の出願人を有し、かつ参照のために本明細書に組み入れられる、米国仮出願第60/347270号（弁理士整理番号021318-000200US）に記載されている。

#### 【0072】

PTSは、メディア・コード変換を実行できるだけでなく、システム・プロトコル・コード変換を実行することもできる。マルチメディア・システム・プロトコルは通常、マルチメ 30  
ディア・エンド・ポイントがどのようにして互いに接続し、コマンドを出し、かつ解釈し（ビデオチャネルのストリーミング、ビデオチャネルの開放など）、接続を切断し、会議に参加することができるかを定義するプロトコルの集合である。システム・プロトコルは通常、呼シグナリング、コマンドおよび制御、メディア・トランスポート局面、ならびにメディア符号化局面のような重要な局面をカバーする。たとえば、H.323システム・プロトコル標準は、呼シグナリングおよびメディア・トランスポートに関するH.225.0/Q.931、コマンドおよび制御に関するH.245、ならびにいくつかのオーディオ・コーデックおよびビデオ・コーデックをカバーする。H.324システム・プロトコル標準は、H.223（メディアおよびデータ・ビットストリーム多重化）、コマンドおよび制御に関するH.245、なら 40  
びにいくつかのオーディオ・コーデックおよびビデオ・コーデックをカバーする。H.323とH.324のいくつかの局面は類似しているが、H.323はパケット・ベースであり、一方、H.324は回線ベースである。したがって、H.323エンド・ポイントとH.324が通信するには、システム・プロトコル・コード変換を呼シグナリング・レベル、コマンドおよび制御レベル、ならびにメディア符号化レベルで行う必要がある。PTSは、システム・プロトコル・コード変換とメディア（オーディオおよびビデオ）コード変換の両方を行う。PTSの観点からすると、呼シグナリング・コード変換は、エンド・ポイント同士が互いに接続できるようにエンド・ポイント（H.324、H.323、SIP、RTSPなど）を代理するプロセスである。コマンドおよび制御の観点からすると、PTSは、端末機能、論理チャネルの開閉などのメッ 50  
ッセージが、コマンドおよび制御メッセージを受信する端末によって理解できるように変換されるようにコード変換を行う。PTSは、送信側エンド・ポイントから受信したメッセー

ジを、受信側端末によって理解できるように変換する必要がある。メディア・トランスポートに関しては、PTSは、回線ペアラ・チャネルを逆多重化し、磁気サービス・データ・ユニットを抽出し、メディア・ビットをコード変換し、次いで、コード変換された各ビットを、受信側エンド・ポイントが理解できるフォーマットで送信できるようにパッケージングすることによって、データにアクセスする必要がある。受信側エンド・ポイントがH.323である場合、パッケージングはRTPパケット化を含む。したがって、メディア・トランスポート・コード変換（変換）は、メディア・ビットの回線パケット、パケット→回線、およびパケット→パケット変換からなる。トランスポートが回線ベースである場合、通常、メディア・ビットは時間ドメイン多重化（TDM）式に多重化される。

【0073】

現在本発明の態様例とみなされているものを図示し説明したが、当業者には、本発明の真の範囲から逸脱せずに、様々な他の修正を加えることができ、均等物を置き換えることができることが理解され则认为される。さらに、本明細書に記載した本発明の中心的概念から逸脱せずに、特定の状況の本発明の教示に適應させるように多数の修正を施すことができる。

【図面の簡単な説明】

【0074】

【図1】 エンド・ユーザと通信する携帯電話の簡略化されたブロック図を示す。

【図2】 代理コード変換サーバの動作の一態様を示す。

【図3】 代理コード変換サーバの他の態様を示す。

【図4】 代理コード変換サーバの接続性の一態様を示す簡略化されたブロック図である。

【図5】 ゲートウェイに接続された代理コード変換サーバの態様を示す簡略化されたブロック図である。

【図6】 コード変換・プロセスの一態様を示す簡略化された流れ図である。

【図7】 メイン・システム・メッセージの一態様を示す簡略化された流れ図である。

【図8】 リソース・メッセージの一態様を示す簡略化された流れ図である。

【図9】 PTS機能・モード・メッセージの一態様を示す簡略化された流れ図である。

【図10】 PTS維持メッセージの一態様を示す簡略化された流れ図である。

【図11】 PTSセッション維持メッセージおよびコード変換・メッセージの一態様を示す簡略化された流れ図である。

【図12】 PTSセッション・レート制御メッセージの一態様を示す簡略化された流れ図である。

【図13】 PTSセッション機能メッセージの一態様を示す簡略化された流れ図である。

【図14】 ネットワーク・アドレス指定メッセージの一態様を示す簡略化された流れ図である。

【図15】 PTSメディア混合メッセージの一態様を示す簡略化された流れ図である。

【図16】 PTS IPRメッセージの一態様を示す簡略化された流れ図である。

【図17】 PTSソフトウェア・モジュールの一態様を示す簡略化された流れ図である。

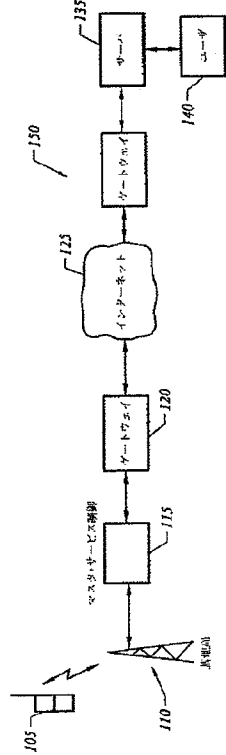
【図18】 各流れ図で使用される記号を示す簡略化された流れ図である。

【図19】 コード変換手順の一態様を示す簡略化された流れ図である。

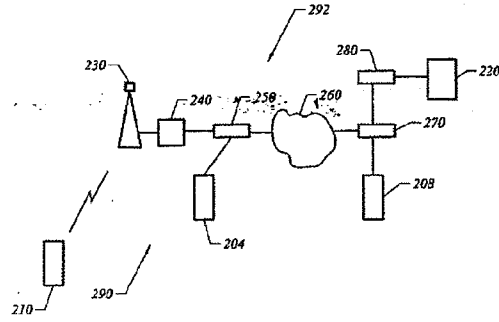
【図20】 PTSのハードウェア・アーキテクチャの一態様を示す簡略化された流れ図である。

【図21】 本発明の態様を実施するのに用いることのできるコンピュータ・システムの簡略化された流れ図である。

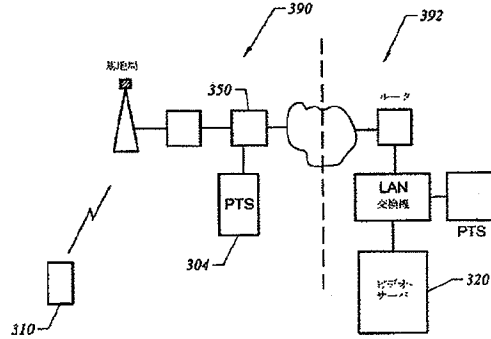
【図 1】



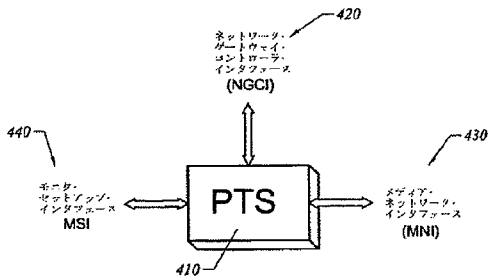
【図 2】



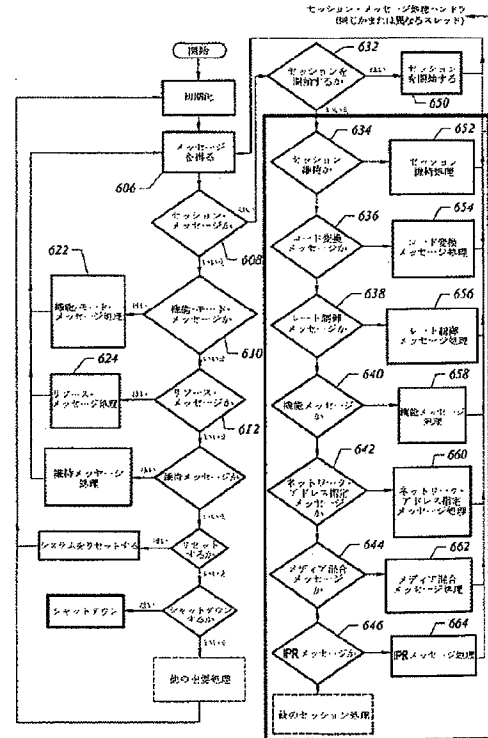
【図 3】



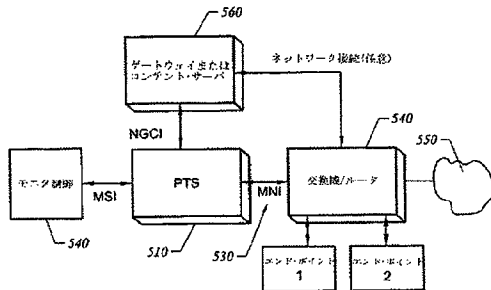
【図 4】



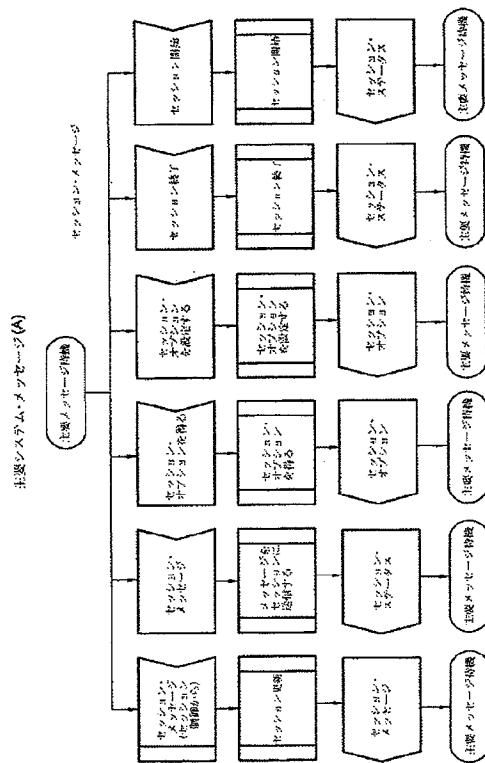
【図 6】



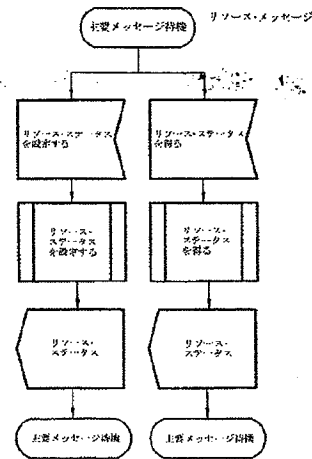
【図 5】



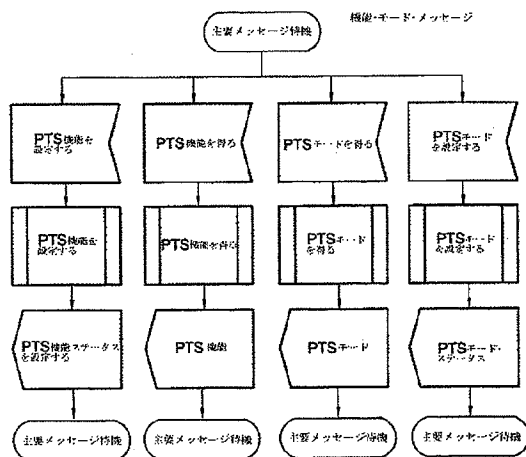
【図 7】



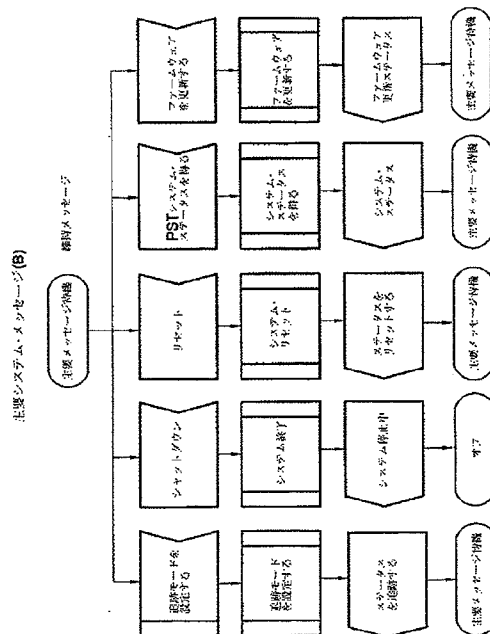
【図 8】



【図 9】

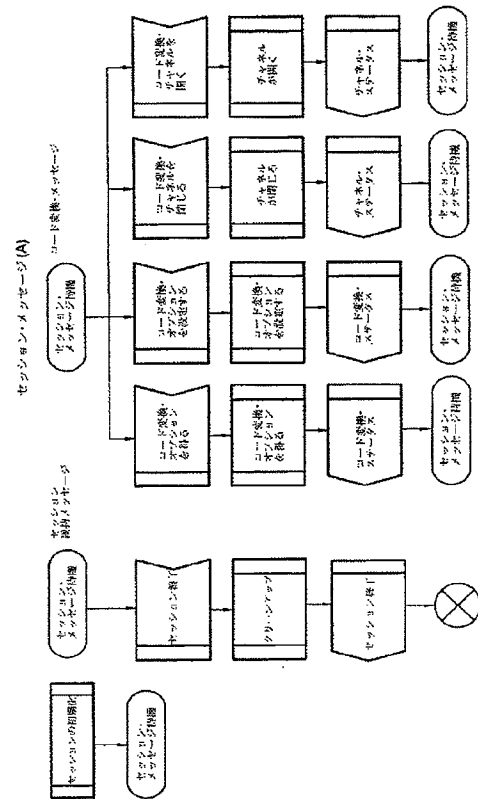


【図 10】

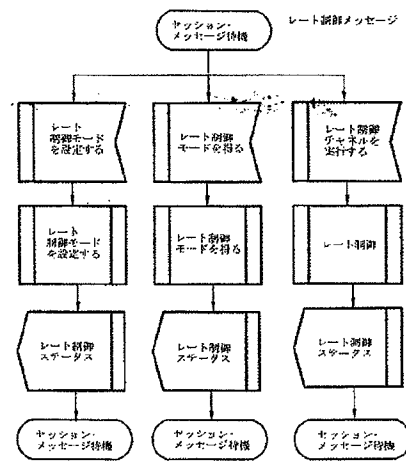




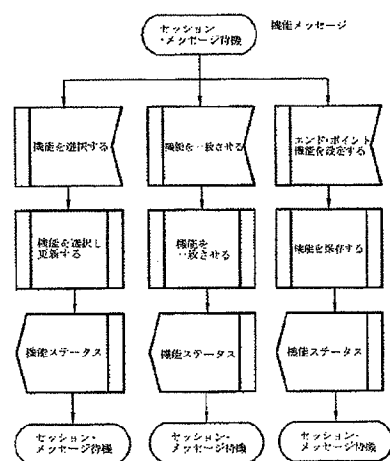
【図 1 1】



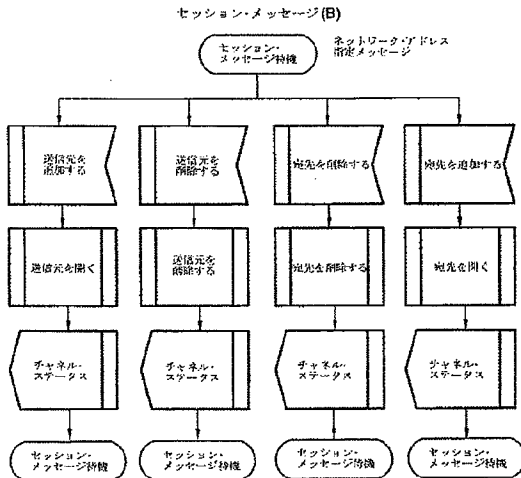
【図 1 2】



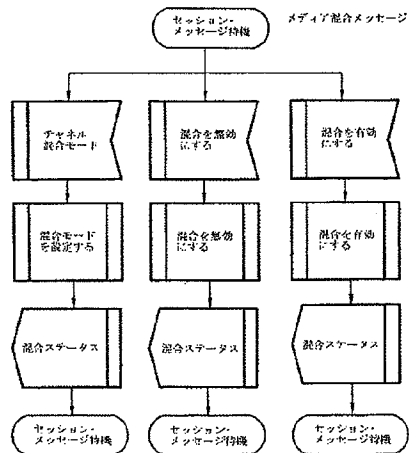
【図 1 3】



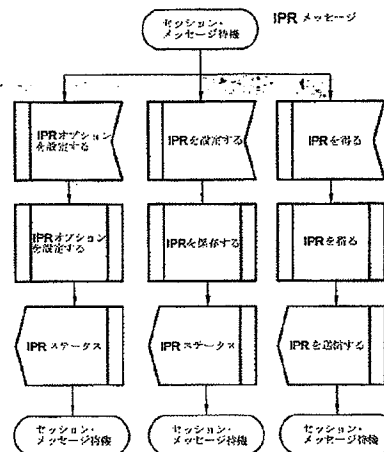
【図 1 4】



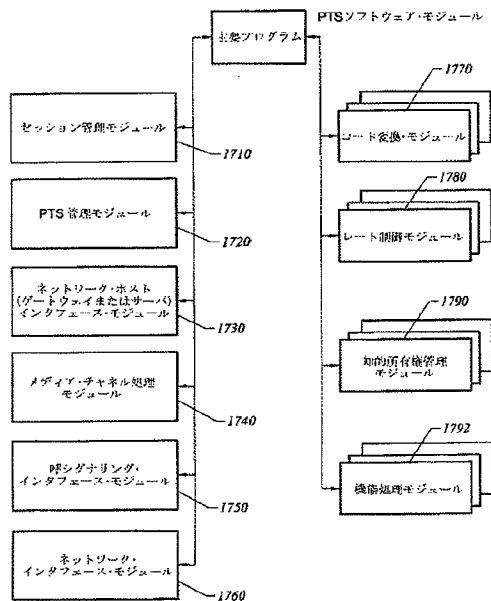
【図 15】



【図 16】

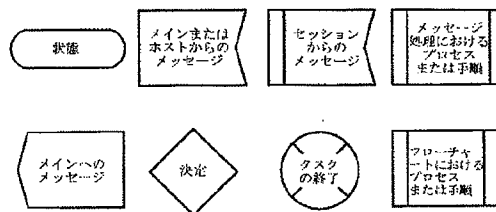


【図 17】



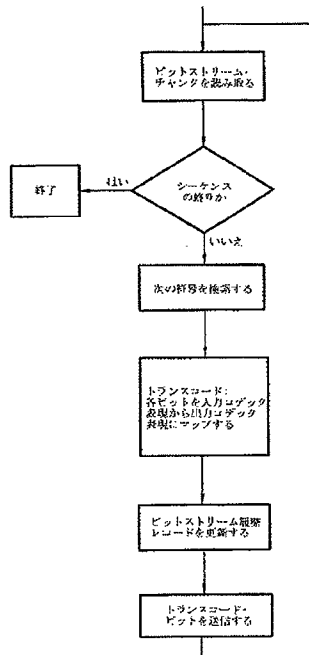
【図 18】

記号

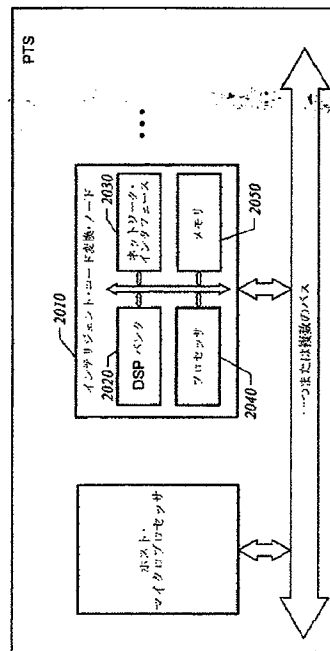


フローチャートおよびメッセージ図で用いられている記号

【図 19】

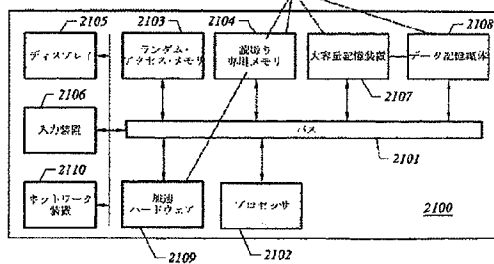


【図 20】



【図 21】

(ファイル内に保存するかまたはメモリに埋め込まれた) PTS ソフトウェア



## 【国際公開パンフレット】

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau(43) International Publication Date  
19 September 2002 (19.09.2002)

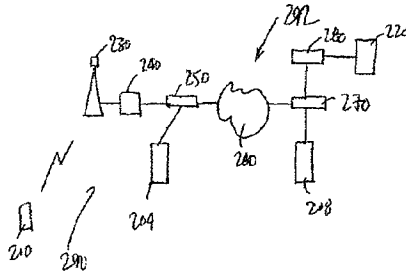
PCT

(10) International Publication Number  
WO 02/073443 A1

- (51) International Patent Classification: G06F 15/16 (74) Agents: OCAWA, Richard, T. et al.: Townsend and Townsend and Crew LLP, 2 Embarcadero Center, 8th Floor, San Francisco, CA 94111 (US).
- (21) International Application Number: PCT/US02/09218
- (22) International Filing Date: 13 March 2002 (13.03.2002)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:  
80/275,584 13 March 2001 (13.03.2001) US  
10009,801 12 March 2002 (12.03.2002) US
- (71) Applicant: DELTATHIN NETWORKS, INC. (US415);  
700 Larkspur Landing Circle, Suite 263, Larkspur, CA  
94939 (US).
- (72) Inventor: JAMRI, Marwan, Awwar, Unit 34, 267 East  
Smyth Street, Sydney, New South Wales 2000 (AU).
- (86) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, FR, GB, GR, GT, HA, HK, HU, IL, IN, IS, JP, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MY, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZM, ZW.
- (86) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW);  
European patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM);  
European patent (AT, BE, CH, CY, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SI, SK, TR);  
OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- Published: with international search report

[Continued on next page]

(54) Title: METHOD AND SYSTEM FOR TRANSCODING VIDEO AND SPEECH SIGNALS



WO 02/073443 A1

(57) Abstract: A system and method for transferring multimedia information from a source location (210) to a destination location (280) through one or more networks (220, 290, 292), which may be different. The system has a source output which provides a first stream of information in a first format. The system also has a destination input which receives a second stream of information in a second format. A proxy transcoding server ("PTS") (240, 294) is coupled between the source output and the destination input. The PTS has a transcoding module transcoding data. The PTS has a transcoding module transcoding data. The PTS also has a capability module identifying a first capability of the source output and a second capability of the destination input, and selecting a transcoding process based upon the first capability and the second capability.

---

**WO 02/073443 A1** 

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

WO 02/073443

PCT/US02/08218

**METHOD AND SYSTEM FOR TRANSCODING VIDEO AND SPEECH  
SIGNALS**

5                   CROSS-REFERENCES TO RELATED APPLICATIONS

[01]   The present application claims priority to U.S. Provisional Patent Application No. 60/275,584 filed March 13, 2001, which is commonly assigned, and hereby incorporated by reference for all purposes.

10                   STATEMENT AS TO RIGHTS TO INVENTIONS MADE UNDER  
FEDERALLY SPONSORED RESEARCH OR DEVELOPMENT

[02]   NOT APPLICABLE

15                   REFERENCE TO A "SEQUENCE LISTING," A TABLE, OR A COMPUTER  
PROGRAM LISTING APPENDIX SUBMITTED ON A COMPACT DISK.

[03]   NOT APPLICABLE

                  BACKGROUND OF THE INVENTION

20           [04]   The present invention is related to the field of communications. In particular, the present invention is related to a method and apparatus for transcoding video and audio signals. More particularly, the invention provides a method and system for transcoding information (e.g., video, voice, data) from a first format to a destination format using a proxy transcoder server having a plurality of transcoding processes, where at least one is selected for  
25 transcoding the information. Merely by way of example, the invention is applied to a wide area telecommunication network, but it would be recognized that the invention can also be applied across many different types of multimedia protocols over transport networks such as the Internet, a mobile network, a local area network, PTSN, ISDN, SONET, DWDM, and others.

                  [05]   Telecommunication techniques have improved dramatically over the  
30 years. Many different types of networks such as fixed switched, packet based, wirelets and

WO 02/073443

PCT/US02/08218

mobile have been deployed. One of the most widely known world wide network called the "Internet" has popularized networking to many people around the world. An increase in use of wide area networks such as the Internet has resulted in many new on-line services such as electronic mail, video telephony, video streaming, electronic commerce, and others. Although computers originally connected to the Internet, other devices such as mobile phones, personal digital assistants, laptop computers, and the like have also been connected. Accordingly, many different types of devices now have access to many different types of services over a variety of networks.

[06] A variety of network elements make up the networks, which connect the aforementioned devices together. Such devices are often connected by gateways and switches that handle transfer of data and conversion of messages from protocols of a sending network to protocols used by a receiving network. Gateways and switches convert analog voice messages to digital formats including G.711 and G.723.1, which are ITU standards. Gateways transmit the converted messages typically in a way similar to transmission of voice over IP. G.711 is an ITU standard for speech codecs that provides audio signals at 64 Kbps using either the A-Law PCM method or the mu-Law PCM method. G.723.1 is an ITU standard for speech codecs optimized for narrow-band networks, including Plain Old Telephone Systems and narrow band Internet connections. The standard uses the LD-CELP method and provides audio signals at 5.3 or 6.3 Kbps. Depending upon the application, there can be many others as well.

[07] As merely an example in Figure 1, a conventional system 100 is shown. This diagram is merely an example and provided for illustrative purposes only. A message originates from a mobile device 105, which is coupled to a wireless network. The message is sent from the mobile device to base station 110 through the wireless network. The base station is coupled to a service station 115, which is coupled to gateway 120. The base station receives the radio message from the mobile device 105, and converts the message, without transcoding, into a digital format, and transmit it to the service station 115. The reformatted message is subsequently transmitted to the gateway which in turn transmits the message to its destination, a user 140, through the Internet 125 and also through a variety of network elements. Such elements may include a gateway 130, a server 135, and others.

[08] One or more gateways may also convert videoconferencing signals from one digital format to another, such as from H.320 to H.323, and transmit converted signals over

WO 02/073443

PCT/US02/08218

the Internet. H.320 is an ITU standard for videoconferencing over digital lines, and it uses the H.261 video compression method, which allows H.320-compliant videoconferencing and desktop systems to communicate with each other over ISDN, switched digital lines and leased lines. H.323 is an ITU standard for real-time, interactive voice and videoconferencing over LANs and the Internet. Widely used for IP telephony, H.323 allows any combination of voice, video and data to be transported. H.323 specifies several video codecs, including H.261 and H.263, and audio codecs, including G.711 and G.723.1. Unfortunately, the audio and video standards have grown well beyond H.320, H.323, G.711, and G.723.1. That is, the proliferation of different standards has caused difficulty in communicating messages between them.

10 Additionally, any communication between such standards has caused a proliferation of complex conversion techniques, which are time consuming and lack efficiency. Accordingly, there is a need for an efficient way to convert information or transcode between various formats in real time. Because some systems such as H.320 and H.324 are circuit switched systems (data is transmitted as a continuous stream of bits) and some other systems are packet based, the connection of circuit-based to packet based systems require the demultiplexing of bits from circuit based bitstreams into packet (circuit-to-packet) and vice versa (packet-to-circuit). Note that different system protocols such as H.320, H.323, H.324, 3GPP-324M, SIP and SDP, make use of different signaling methods (to setup connections and exchange terminal capabilities). The inter-connectivity of these systems require the trans-signaling and the conversion of terminal capabilities so terminal can understand what terminals using different protocols are capable of.

20

[09] From the above, it is seen that an improved way of transferring information from a source to a destination is highly desirable.



WO 02/073443

PCT/US02/08218

## SUMMARY OF THE INVENTION

[10] According to the present invention, improved techniques for transcoding in the telecommunication fields are provided. In particular, the present invention is related to a method and apparatus for transcoding video and speech signals. More particularly, the invention provides a method and system for transcoding information (e.g., video, voice, data) from a first format to a destination format using a proxy transcoder server having a plurality of transcoding processes, where at least one is selected for transcoding the information. Merely by way of example, the invention is applied to a wide area telecommunication network, but it would be recognized that the invention can also be applied across many different types of multimedia protocols over transport networks such as the Internet, a mobile network, a local area network, PTSN, ISDN, SONET, DWDM, and others.

[11] In a specific embodiment, the invention provides a system for transferring multimedia information from a source location to a destination location through one or more networks, which may be different. The system has a source output which provides a first stream of information in a first format. The system also has a destination input which receives a second stream of information in a second format. A proxy transcoder server ("PTS") is coupled between the source output and the destination input. The PTS has a transcoding module transcoding data. The PTS also has a capability module identifying a first capability of the source output and a second capability of the destination input, and selecting a transcoding process based upon the first capability and the second capability. Preferably, the selecting is provided using capability mode selection.

[12] In an alternative specific embodiment, the invention provides a system for transferring multimedia information from source to destination locations through one or more networks, which may be different. The system has a source output in a first format, where the source output is coupled to a first network, the source output providing a first stream of information. The system also has a destination input to be received in a second format, where the destination input is coupled to a second network. The destination input receives a second stream of information. A proxy transcoder server ("PTS") is coupled between the source output and the destination input. The proxy transcoder server has a capability process, which is adapted to identify a first capability of the source terminal (which may have different capabilities) and is

WO 02/073443

PCT/US02/08218

adapted to identify a second capability of the destination terminal (which may also have different capabilities). The server also has a transcoding process comprising a plurality of transcoding modules numbered 1 through N, where N is an integer greater than 1. The transcoding process is adapted to selected one of the transcoding process based upon the first capability and the second capability. The proxy transcoder server has a bit rate control process. The bit rate control process is adapted to receive a network status information (e.g., ping) from the first network. The bit rate control is adapted to adjust a status (e.g., stop, prioritize allow, adjust bit rate (by selecting lower bit rate coder)) of the stream of information based upon the network status information.

10 [13] In an alternative specific embodiment, the invention provides a method for processing streams of information. The method includes identifying a source capability from a plurality of source capabilities for a stream of information. The method also identifies a destination capability from a plurality of destination capabilities. A step of selecting a transcoding process from a plurality of transcoding processes in a library based upon the identified source capability and the identified destination capability is included. The method also processes the stream of information using the selected transcoding process if the identified source capability and the identified destination capability are different. The method also transferring the stream of information from the source to the destination free from one of the transcoding processes if the identified source capability and the identified destination capability matches.

20 [14] Numerous benefits are achieved using the present invention over conventional techniques. In a specific embodiment, the invention provides a way to transcode video data from H.263 to MPEG-4 data and vice versa (and between other video codecs), or to transcode audio data from G.723.1 to GSM-AMR (and between other audio codecs) and vice versa. Transcoding can proceed seamlessly so that end points receiving transcoded data do not notice the conversion in preferred embodiments. The invention can also be implemented using conventional software and hardware technologies, such as digital signal processors (DSPs). Depending upon the embodiment, one or more of these benefits or features can be achieved. These and other benefits are described throughout the present specification and more particularly below.

WO 02/073443

PCT/US02/08218

[15] The accompanying drawings, which are incorporated in and form part of the specification, illustrate embodiments of the invention and, together with the description, serves to explain the principles of the invention.

## BRIEF DESCRIPTION OF THE DRAWINGS

- 5 [16] Figure 1 illustrates a simplified block diagram of a mobile phone communicating with an end-user.
- [17] Figure 2 illustrates one embodiment of the operation of proxy transcoder servers.
- [18] Figure 3 illustrates another embodiment of the operation of proxy transcoder servers.
- 10 [19] Figure 4 is a simplified block diagram illustrating one embodiment of the connectivity of the proxy transcoder server.
- [20] Figure 5 is a simplified block diagram illustrating an embodiment of a proxy transcoder server connected to a gateway.
- 15 [21] Figure 6 is a simplified flow diagram illustrating one embodiment of the transcoding process.
- [22] Figure 7 is a simplified flow diagram illustrating one embodiment of the main system messages.
- [23] Figure 8 is a simplified flow diagram illustrating one embodiment of the resource messages.
- 20 [24] Figure 9 is a simplified flow diagram illustrating one embodiment of the FTS features and mode messages.
- [25] Figure 10 is a simplified flow diagram illustrating one embodiment of the FTS maintenance messages.
- 25 [26] Figure 11 is a simplified flow diagram illustrating one embodiment of the FTS session maintenance and transcoding messages.
- [27] Figure 12 is a simplified flow diagram illustrating one embodiment of the FTS session rate control messages.
- [28] Figure 13 is a simplified flow diagram illustrating one embodiment of the
- 30 FTS session capability messages.

WO 02/073443

PCT/US02/08218

[29] Figure 14 is a simplified flow diagram illustrating one embodiment of the network addressing messages.

[30] Figure 15 is a simplified flow diagram illustrating one embodiment of the PTS media mixing messages.

5 [31] Figure 16 is a simplified flow diagram illustrating one embodiment of the PTS IPR messages.

[32] Figure 17 is a simplified flow diagram illustrating one embodiment of the PTS software modules.

10 [33] Figure 18 is a simplified flow diagram illustrating symbols used in the flow diagrams.

[34] Figure 19 is a simplified flow diagram illustrating one embodiment of the transcoding procedure.

[35] Figure 20 is a simplified block diagram illustrating one embodiment of a PTS's hardware architecture.

15 [36] Figure 21 is a simplified block diagram of a computer system that may be used to implement an embodiment of the invention.

#### DESCRIPTION OF THE SPECIFIC EMBODIMENTS

20 [37] According to the present invention, improved techniques for transcoding in the telecommunication fields are provided. In particular, the present invention is related to a method and apparatus for transcoding video and audio signals. More particularly, the invention provides a method and system for transcoding information (e.g., video, audio, data) from a first format to a destination format using a proxy transcoder server having a plurality of transcoding processes, where at least one is selected for transcoding the information. Merely by way of  
25 example, the invention is applied to a wide area telecommunication network, but it would be recognized that the invention can also be applied across many different types of multimedia protocols over transport networks such as the Internet, a mobile network, a local area network, PTSN, ISDN, SONET, DWDM, and others.

30 [38] Parts of the description will be presented in terms of operations performed through the execution of programming instructions according to various embodiments. As will be understood by those skilled in the art, these operations often take the form of electrical,

WO 02/973443

PCT/US02/08218

magnetic, or optical signals capable of being stored, transferred, combined, and otherwise manipulated through, for instance, electrical components. Parts of the description will be presented using a distributed computing environment. In a distributed computing environment, file servers, computer servers, and memory storage devices may be located in different places, but they are accessible to local processing units through the network. In addition, program modules may be physically located in different local and remote memory storage devices. Execution of the program modules may occur locally in a stand-alone manner or remotely in a client-server manner. Examples of such distributed computing environments include local area networks of an office, enterprise-wide computer networks, and the global Internet.

[39] Additionally, the following terms are provided to assist the reader in explaining aspects of the invention. Such terms are not intended to be limiting but merely provided for descriptive purposes to one of ordinary skill in the art. Other meanings for the terms consistent with those understood by one of ordinary skill in the art may also be used.

Term	Description
ASIC	Application Specific Integrated Circuit
CIF	Common Intermediate Format
ETSI	European Telecommunications Standards Institute
G.723.1	ITU Recommendation G.723.1, Dual rate speech coder for multimedia communications transmitting at 5.3 and 6.3 kbit/s, 1996
GOB	Group of blocks
GSM	Global System for Mobile communications
GSM-AMR	ETSI Adaptive Multi Rate speech coder, GSM 06.90: "Digital cellular telecommunications system (Phase 2+): AMR speech transcoding", 1998
GSM-AMR	GSM- Adaptive Multi Rate
H.320	ITU Recommendation H.320, Narrow-band visual telephone systems and terminal equipment, 1997

WO 02/073443

PCT/US02/08218

H.323	ITU Recommendation H.323, Packet-based multimedia communications systems, 1998
H.324	ITU Recommendation H.323, Terminal for low bit-rate multimedia communication, 1998
H.261	ITU Recommendation H.263, Video codec for audiovisual services at $p \times 64$ kbit/s, 1993
H.263	ITU Recommendation H.263, Video coding for low bit rate communication, 1998
IETF	Internet Engineering Task Force
ISO	International Standard Organisation
ITU	International Telecommunication Union
MB	Macro block
MPEG	Moving Picture Expert Group, part of the International Standard Organisation
MPEG2	MPEG audiovisual standards 13818 series
MPEG4	MPEG audiovisual standards 14496 (1-5)
MVD	Motion Vector Data
P Frame or P Picture	Video frame based on predicted information
PTS	Proxy Transcoder Server
QCIF	Quarter CIF (see CIF)
RFC	Request for Comment
SDP	Session Description Protocol
SIP	Session Initiation Protocol
TCOEF or TCOEFF	Transform Coefficients
W3C	World Wide Web Consortium
WAP	Wireless Access Protocol

[40] In a specific embodiment, the term "Proxy Transcoder Server" (herein PTS) is a computer having various modules to carryout some or all of the functionality described herein as well as other functionality, which are known or not known at this present time. The

WO 02/073443

PCT/US02/06218

PTS comprises a host processor, one or more network interfaces, and one or more transcoders. A transcoder may contain, among others, a printed circuit board, an application specific integrated circuit (ASIC), and a field programmable gate array (FPGA). The PTS may be connected to a network host such as a media gateway controller or a soft switch or to a content server such as a video server. The PTS can provide codecs and interfaces to various network architectures and protocols, including WAN, LAN, Mobile, PTSN, ISDN, SONET, and perform one or more of the following functions:

- [41] 1. Capability matching and mode selection;
- [42] 2. Media bitstream transcoding;
- 10 [43] 3. Media bitstream rate control;
- [44] 4. Intellectual property rights management and processing;
- [45] 5. Audio data mixing; and
- [46] 6. Encryption and/or decryption
- [47] The overall function of the PTS is to translate between various protocols,
- 15 as exemplified below.
- [48] 1. Translating multimedia systems protocols, including the ITU H.32X series, e.g., H.242 and H.245, which are used for conveying the media capabilities of end-points, and managing media channels and conferences.
- [49] 2. Translating media streams including audio and video streams, e.g.,
- 20 translating video streams to and from any pair of video codecs including MPEG2, MPEG4, H.261, H.263, or audio streams to and from any pair of audio codecs including G.723.1, G.729, GSM-AMR, EVRC, SMV, and QCELP.
- [50] 3. Translating IP management protocols and IP rights markings. For example, the PTS may extract IP rights information from MPEG4 streams and process the
- 25 information in accordance with IP rights-related operations in order to maintain the IP rights in the media.
- [51] 4. Encrypting and/or decrypting signals if necessary.
- [52] Hence, the PTS has a variety of audio and video transcoding capabilities. Selection of capabilities affects overall qualities of service (QoS) in multimedia communication,
- 30 and should therefore be determined based on the connection bandwidth and load of the gateway to which the PTS is coupled. The PTS can transcode for different media contents including

WO 02/073443

PCT/US02/06218

MPEG series, H.26X video series, GSM-AMR, and G.72X audio codec series. Furthermore, the PTS' transcoding capabilities can be easily upgraded, at least in part due to its programmability.

5 [53] In addition, the PTS can perform media bitstream rate control. The rate control is necessary because the connection between two end-points may involve bandwidth reduction, e.g., from wired to wireless network. The PTS may perform rate control using information from network channel allocation as supplied by the network access provider, or using in-band bandwidth management requests, which can be originated from end-points through command and control protocols, e.g., H.242 and H.245.

10 [54] Moreover, the PTS can perform intellectual property (IP) rights management and processing. For example, the PTS may identify a data set regarding intellectual property rights, and use it to facilitate the management and processing. In one embodiment, MPEG4 supplements coded media objects with an optional intellectual property identification (IP) data set, which carries information on contents, types of contents, and IP rights holders. 15 The data set, if present, forms part of an elementary stream descriptor describing streaming data associated with a media object. The number of data sets associated with each media object may vary; and different media objects may share the same data set. The provision of the data sets allows the implementation of mechanisms for audit trail, monitoring, billing, and copy protection.

20 [55] Multimedia communication applications have wide ranges of requirements for IP rights protection and security. Some applications require protecting information that users exchange in order to preserve privacy, even if the information has no intrinsic value. Other applications require high-grade management and protection for information of importance to its creator and/or distributors. In addition, the framework of IP rights management and processing 25 must generally be flexible so as to access various forms of IP rights data, as required by a specific application, and as stored in a particular bitstream.

[56] Although the description that follows shows a PTS connected to a gateway as a stand alone device, alternate embodiments of the PTS may use a PTS connected to a switch, a server, a router or any device connected to a network. In addition, a PTS may be integrated 30 into a gateway, a switch, a router or other device connected to a network to form an integral part of that network device. Further details of other embodiments are provided below.



WO 02/073443

PCT/US02/09218

[57] Figure 2 illustrates an embodiment wherein a PTS transcodes signals between a mobile end-point and a LAN phone. This diagram is merely an example, which should not unduly limit the scope of the claims herein. One of ordinary skill in the art would recognize many other variations, modifications, and alternatives. A mobile end-point 210 communicates audio and video information with a LAN phone 220. The transmitted information travels from the mobile end-point 210 to a Base Station 230 via a radio communication channel, e.g., an air link, and subsequently to a Master Service Controller 240, to a near-end gateway 250, through the Internet 260, to a far-end gateway 270, to a LAN switch 280, and finally to the LAN phone 220.

[58] Because of limitations of an air-link and mobility of the mobile end-point 210, the bandwidth between the mobile end-point 210 and the base station 230 may reach several tens or hundreds of Kbps, much smaller than the bandwidth between Internet routers and the LAN phone 220. The latter bandwidth reaches 10s to 100s of megabits per second (Mbps). In Figure 2, the media coding and decoding capability of the mobile end-point 220 is GSM-AMR for audio signals and MPEG-4 for video signals. In contrast, the media coding and decoding capability of the LAN phone 230 is G.723.1 for audio and H.263 for video. The differences in capability and bandwidth make important transcoding performed by PTS in the path between the mobile end-point 220 and the LAN phone 230, to facilitate substantially real-time communications. In addition, transcoding by PTS prevents overloading a mobile network 290, because the PTS performs a throttling function.

[59] In Figure 2, audio signals need to be transcoded between G.723.1 and GSM-AMR, while video signals need to be transcoded between MPEG4-video and H.263. Such transcoding may be performed by either a PTS 204 in the mobile network 290 or a PTS 208 in a landline network 292. Thus the two PTS's 204 and 208 are not necessary to perform the transcoding. But implementing one PTS in each network allows the bandwidth requirement at each gateway to become smaller. In addition, even though the embodiment illustrated in Figure 2 shows transcoding by PTS, the PTS's can also allow transmission of signals without transcoding.

[60] Figure 3 illustrates a block diagram of yet another embodiment of the present invention. This diagram is merely an example, which should not unduly limit the scope of the claims herein. One of ordinary skill in the art would recognize many other variations,

WO 02/073443

PCT/US02/08218

modifications, and alternatives. A mobile handset 310 communicates information with a video content server 320. The video server 320 streams movies containing audio and video signals. The audio signals are encoded using MPEG2-Audio Level 3 (MP3) and the video signals are encoded using MPEG2-video. Due to the mismatch in bandwidth and capability between the video server 320 and the mobile handset 310, the PST 304 in a mobile network 390 would transcode the audio to GSM-AMR and the video to MPEG4-video. But if the PST 308 in a video server network 392 performs the transcoding, the bandwidth requirement for a gateway 350 in the mobile handset network 390 may be reduced.

[61] In yet another embodiment, the capabilities of two end-points of communication may be specified using Wireless Access Protocol (WAP) User Agent Profile information (Capabilities and Preferences Information), the web consortium CC/PP (using the Resource Description Framework, or RDF), the IETF standards (RFC 2506, RFC 2533, and RFC 2703) or ITU's H.245 or H.242 standards, or a combination thereof. In yet another embodiment, a gateway may, under the assistance of PTS, detect the capabilities of end-points and the bandwidth available or allocated for transmission of media. The PTS may then select a data coding mode and transcode the media to best meet the requirements of the end-points. Hence, the PTS may serve as a proxy for each end-point side, converting bitstreams in one form appropriate for one side to another form appropriate for the other side.

[62] Figure 4 illustrates a block diagram of an PTS according to an embodiment of the present invention. This diagram is merely an example, which should not unduly limit the scope of the claims herein. One of ordinary skill in the art would recognize many other variations, modifications, and alternatives. Logical ports are defined as abstract communication ports. Several logical ports may exist on one physical network interface, but a PTS may have more than one physical network interface. A network gateway controller interface (NGCI) 420 may comprise at least one NGCI logical port for messaging between a PTS 410 and a gateway or a network host. The gateway may perform call signaling between end-points. The gateway may be a media gateway controller, or any other gateway device that establishes call or transport function between end-points. Call signaling is the process of establishing the initial link between end-point entities, by exchanging network addresses of the source and the destination of a media stream to be transcoded. For real-time Internet protocol, call signaling establishes Internet protocol addresses and Internet protocol port numbers for the

WO 02/073443

PCT/US02/08218

source and the destination of a media stream. Call signaling could also involve more sophisticated processes, such as that required by SIP or H.323.

{63} Media network interfaces (MNI) 430 provide logical ports for reception and transmission of media bitstreams. Through these logical ports, the PTS 410 receives media bitstreams, e.g., audio signals, video signals, command and control data, and other data that could be text or binary. Port addresses for receiving bitstreams by the PTS 410 and the destination addresses to which the transcoded streams are transmitted are specified by a messaging protocol between the call signaling gateway and the PTS 410 through the NGCI 420. The underlying physical interface of the MNI 430 could be a Gigabit network interface card, time domain multiplexed (TDM) circuit-switched connections such as E1/T1/OC3. The MNI 430 physical network interfaces may or may not share the physical link of the NGCI 420.

{64} A monitor and setup interface (MSI) 440 is used for initial configuration, monitoring, and reconfiguration of the PTS 410. The physical link over which the MSI data are carried may be shared with other network connections or may be carried through a dedicated serial connection port. Many connectivity scenarios exist for a PTS, including connecting to a gateway e.g., a media gateway controller, and connecting to a content server, e.g., a video server.

{65} Figure 5 illustrates the connectivity of a PTS according to an embodiment of the present invention. This diagram is merely an example, which should not unduly limit the scope of the claims herein. One of ordinary skill in the art would recognize many other variations, modifications, and alternatives. A MNI 530 connects a PTS 510 to the Internet 550, via a router or switch 540, and a NGCI connects the PTS 510 to a content server (or gateway 560 can be a gateway mediating between two end-points, e.g. a mobile and an IP phone) 560 via the NGCI 520. The MSI 530 may also in connection to the content server 560 if the content server can provide basic terminal emulation support, e.g., hyper-terms under Windows. In one embodiment, the gateway or content server may be a network host.

{66} Figure 6 illustrates various PTS functions the means of an operation cycle according to an embodiment of the present invention. This diagram is merely an example, which should not unduly limit the scope of the claims herein. One of ordinary skill in the art would recognize many other variations, modifications, and alternatives. Once a PTS is connected to the network, e.g., a network host or a router, and is powered-up and configured, the PTS is ready to accept the initiation of transcoding sessions from the network host.

WO 02/073443

PCT/US02/06218

[67] At step 606, the PTS receives a message from either a gateway or from a content server. At step 608, the PTS determines if the received message is a session message. If the message is a session message, at step 632, the PTS decides if the message is a start of session message. A session may be initiated by a network host through a messaging protocol operated over the NGCI, such as the MGCP or the H.248/MEGACO protocols. One or more sessions can be active at any time. Once a session is started, basic operations of the PTS for that session are controlled by further messages.

[68] If the message received is a start of session message, at step 650, the PTS starts a new session. If the received message is not a start of session message, at step 634, the PTS decides whether the received message is a session maintenance message. If it is a session maintenance message, the PTS commences a session maintenance process at step 652. The session maintenance process may involve maintaining or terminating a connection for a session. If the received message is not a session maintenance message, at step 636, the PTS decides whether the received message is a transcoding message. If the message is a transcoding message, at step 654, the PTS processes the transcoding message. If the message is not a transcoding message, at step 638, the PTS determines whether the received message is a rate control message. If the received message is a rate control message, at step 656, the rate control message is processed. The rate control process may direct the PTS to dynamically adjust the transmission rate in order to efficiently utilize the transmission bandwidth and to prevent transmitted packets from being dropped.

[69] When the PTS transcodes a bitstream, the optimal data rate produced by the PTS depends on media protocol of the network between the PTS and the destination of the bitstream. The data rate received by the destination varies, depending on network congestion, type of link between a router and the destination, e.g., wired connection or wireless connection, protocol and data multiplexing associated to the link, and the quality of the link. The destination may receive a data rate higher than it can process, resulting in buffer-overflow. On the other hand, the destination may receive a data rate lower than the rate it expects to receive, resulting in buffer-underflow. Therefore, the PTS may have to regulate its data rate to avoid both buffer-overflow and buffer-underflow.

[70] The PTS regulates its data rate with at least the following methods. The PTS may use network congestion information, bandwidth information, quality information from

WO 02/973443

PCT/US0208218

the network host or the network access provider, or from internal PTS mechanisms to calculate roundtrip time. The roundtrip time between the PTS and an end-point, e.g., source or destination of a bitstream, can be measured by sending a "ping" packet to the end-point. The total time for the "ping" packet to reach the end-point and for the response packet from the end-point to reach the PTS is the roundtrip time. The more congested the network is, the longer is the roundtrip time. Hence the roundtrip time can be used to assess the congestion level of the network at current PTS bit-rate.

[71] Alternately, the PTS may assess the network congestion using in-band information. For example, the PTS may receive from an end-point, instructions to reduce or to increase its bit-rate throughput, under protocols such as H.245 used in H.324 and H.323.

[72] Either the in-band method or the roundtrip-time method may be used to maintain adequate quality of service given instantaneous network conditions. The PTS uses the congestion and bit-rate information to determine an adequate coding mode that yields an appropriate bit-rate with the following methods.

[73] In one embodiment, the PTS may change transcoding parameters to satisfy service goals in real-time operation. For example, in MPEG4-video, the quantization parameters can be changed to yield a low bit-rate throughput. But the change may lead to degradation in video quality. Therefore if video quality is important, advanced coding techniques such as that provided by H.263 and MPEG4 may be used to reduce the bit-rate throughput without degrading quality. But these methods may impose higher computation requirements. Hence, the PTS rate control strategy should provide a desired balance between signal quality, bit-rate and computation.

[74] The rate control in audio is achieved similarly although most audio codecs cannot provide fine-grained variable rates, but instead provide a number of bit rates that the PTS can select from. For example, the G.723.1 audio codec provides two bit-rates, a low rate and a high rate. Similarly, the GSM-AMR codec supports eight bit-rates, ranging from 4.75 Kbps to 12.2 Kbps. The PTS may use a lower rate if the network path to the end-point has congestions, or if the bandwidth allocated to the link with the end-point has a low bandwidth.

[75] In yet another embodiment a PTS may instruct the network equipment such as a router to give a higher priority to the data being handled by the PTS, if the network equipment supports prioritization of data delivery. For instance, version 6 of the Internet

WO 02/073443

PCT/US02/08218

protocol provides support for prioritization of packets. Also, the IETF has developed standards for resource reservation that allows end-points to reserve bandwidth. The PTS may exploit this packet prioritization and resource reservation if they are supported by network in which the PTS is being deployed. For example, the Internet protocol provides facility for assigning priority to packets, and the PTS can use this facility to prioritize packets when needed. In addition, the PTS may support internal prioritization of its own connections so connections of higher priority are processed as soon as they become practical.

[76] Turning back to Figure 6, at step 640, the PTS determines whether the message received is a capability message. A capability message is a message that contains capabilities of the end-points. If the message is a capability message, at step 658, the PTS processes this message. The PTS can process the message defining capabilities of end-points. For example, if the end-points are a video server and a mobile terminal, the capabilities of the video server and the mobile terminal are conveyed to the PTS. Consequently, the PTS determines the best mode of communication between these two end-points. In choosing the best mode, the PTS takes into account the protocol associated with the specific video content that the mobile user would like to view. Different types of contents may be encoded with different protocols, e.g., MPEG2 and MPEG4. In addition, the capabilities of the mobile terminal may be communicated to the PTS via the network host in numerous ways. For example, the PTS may obtain the capabilities from information stored in the mobile terminal, from the user subscription information stored in the network database of the user's service provider, or in-band within the bitstream exchanged between the mobile terminal and the network access gateway during the call signaling phase. The format of the capabilities includes among others ITU, IETF, and WAP.

[77] A capability message may be sent by the network host to the PTS to find the best transcoding mode for a specific media to be transmitted from one end-point, e.g., a video server, to the other end-point. In the capability-mode-selection process, the PTS may choose one bitstream protocol mode for receiving data from the source, and another bitstream protocol mode to which the PTS convert the received media.

[78] In yet another embodiment, the selected modes for each end-point may be signaled to the respective end-point in order to open bitstream transport channels. In H.323 or H.324, the network uses H.245 logical channel operations or Fast Connect procedures to open such channels. Using H.245 logical channel operations, an end-point may send an "open logical

WO 02/973443

PCT/US02/48218

channel" request to the other end-point in order to transmit signals. In H.323, an end-point can encapsulate information about media channels through which the end-point is ready to transmit signals by using a "fast start" message encapsulated in the call signaling information, exchanged in the initial setup of the call under, for example, the ITU Q.931 standard as recommended by H.225.0. Therefore, given the selected protocols for the end-points, the network host or the PTS can establish transport channels for a media bitstream. The opening process depends on the system level protocol of the overall connection between the end-points mediated by the network.

[79] Before the PTS can transcode, the selected media transmission modes have to be associated with a source address and a destination address, and the information of such association has to be communicated to the PTS via messages. The association may result from opening a logical media channel by one of the end-points as done in the H.245 standard. This implicit opening of a logical media channel when the PTS selects the types of media transmission modes can be explicitly requested by the gateway or the content server, or could be pre-programmed under some standard requirements. Regardless of the means for specifying the source address or the destination address, the association between selected media transmission modes and the source or destination addresses informs the PTS where to get the input bitstream and where to send the transcoded bitstream.

[80] In a specific embodiment, the PTS reads the bitstream from the source address, translates the bitstream from its original format to the target format, and sends the converted bitstream to the destination address. The reception and transmission of bitstream data is performed with network read/write functions, using the network hardware-specific software.

[81] Turning back to Figure 6, at step 642, the PTS determines whether the message received is a network-addressing message. A network-addressing message contains information about network addresses for the source and/or the destination of a bitstream. If the message received is a network-addressing message, at step 660, the PTS processes the network-addressing message. At step 644, the PTS determines whether the message received is a media-mixing message. A media-mixing message is a message requesting the PTS to mix signals associated with two or more audio streams and retransmit the mixed bitstreams to a network destination address. If the received message is a media-mixing message, at step 662, the PTS processes the media-mixing message. At step 646, the PTS determines whether the message received is a IP rights message, and if so at step 664 the PTS manages the information regarding

WO 02/073443

PCT/US02/08218

IP rights contained in the received message, according to instructions in the message. Some media communication and representation protocols support interfacing with IP rights management and processing, and accessing information regarding IP rights contained in bitstreams. For example, the PTS may support the MPEG-4 interface specification for IP rights management and processing. The information regarding IP rights is extracted or demultiplexed from MPEG4 bitstreams and made available to the network host via the messaging system. The IP rights specific applications installed on the network host or plugged into the PTS then access and process such information for various purposes, including record keeping, resigning of the content, and blocking.

[82] At step 610 in Figure 6, the PTS determines whether the received message is a features and mode message and if so, at step 622, the PTS processes this message. Processing a features and mode message includes activating features and options that are related to transcoding, mixing and other options associated to the session. At step 612, the PTS determines whether the received message is a resource message, and if so, at step 624, the PTS processes this message. Processing a resource message includes processing the PTS resources in terms of hardware processing resources, memory resources, and other computing or networking resources that the PTS manages. If the PTS receives instructions that the media being received and thereafter transmitted needs to be encrypted, the PTS transcodes and thereafter encrypts the transmitted data. Similarly, the PTS may decrypt data in response to instructions received.

[83] In the above, the capabilities of the PTS are illustrated as discrete steps, but one skilled in the art will appreciate that one or more of the discrete steps may be combined or further subdivided to perform functions of the discrete steps. Depending upon the embodiment, the functionality described can be separated or even combined. The functionality can be implemented in software and/or hardware including any combination of them. Depending upon the embodiment, there can be many other modifications, variations, and alternatives.

[84] Figures 7, 8, 9, and 10 are simplified flow diagrams illustrating an embodiment of the main system messages. These diagrams are merely examples, which should not unduly limit the scope of the claims herein. One of ordinary skill in the art would recognize many other variations, modifications, and alternatives. The main system messages allows the PTS to respond to instructions from a network equipment such as a media gateway controller. The messages includes instructions to start a session, end a session, set session options, get



WO 02/073443

PCT/US82/09218

session options, send a message to session manager, get message from session manager, set PTS mode, get PTS mode, set PTS feature, get PTS feature, get resource status, set resource status, update firmware procedure, get PTS system status, reset the PTS, shutdown the PTS, and activating debug/tracing mode.

5 [85] Figures 11-16 are simplified flow diagrams illustrating an embodiment of the PTS session maintenance and transcoding messages. These diagrams are merely examples, which should not unduly limit the scope of the claims herein. One of ordinary skill in the art would recognize many other variations, modifications, and alternatives. The session maintenance and transcoding messages allow the PTS to terminate a session, open/close transcoding channel, 10 set/get transcoding options, updating end-point capabilities, match end-points capabilities, select end-point capabilities, activate rate control, get rate control mode, get rate control mode, add/remove media destination addresses, add/remove media source addresses, set/get IPR modes, set IPR options, activate mixing of media channels, disable mixing, and set channel mixing mode.

15 [86] Figure 17 shows a block diagram illustrating software modules that may be used in a PTS according to an embodiment of the present invention. This diagram is merely an example, which should not unduly limit the scope of the claims herein. One of ordinary skill in the art would recognize many other variations, modifications, and alternatives. The PTS software comprises the following main modules:

- 20 [87] 1. A session management module 1710;  
 [88] 2. A PTS management module 1720;  
 [89] 3. A network host interface module 1730;  
 [90] 4. A media channel processing module 1740;  
 [91] 5. A call signaling interface module 1750;  
 25 [92] 6. A network interface module 1760;  
 [93] 7. A transcoding module 1770;  
 [94] 8. A rate control module 1780;  
 [95] 9. An intellectual property rights management module 1790; and  
 [96] 10. A capabilities processing module 1792.  
 30 [97] The session management module 1710 performs the main services of a gateway, and is thus the main PTS software program. For example, the session management

WO 02/073443

PCT/US02/08218

module 1710 starts and ends transcoding sessions, handles and dispatches session messages, and manages session resources. The PTS management module 1720 performs basic overall management functions necessary if an operator needs to inspect the status of the PTS or to manage its resources. For example, the management module 1720 tests main hardware components with component specific testing procedures, resets the PTS, and tracks and allocates transcoding resources dynamically. The network host interface module 1730 handles communication-messaging interface between the PTS and the network host, e.g., a media gateway controller or a content server. For example, the interface module 1730, depending on type of network host, implements messaging between the network host and the PTS, implements commands defining methods by which transcoding capabilities are captured by the PTS or defined for the PTS. In addition, the interface module 1730 may implement methods by which the PTS retrieves media content types from bitstreams when capability exchanges are not explicitly performed. The media channel processing module 1740 performs media channel functions, such as opening, closing, adding and removing channel network sources and destinations. The call signaling interface module 1750 performs functions for establishing initial call setup between end-points through the PTS, where the procedure for call setup depends on standard, such as SIP and Q.931. The network interface module 1760 provides basic input and/or output communication interfaces. The basic input and/or output is the lowest level of communication over which more complex messaging is performed.

[98] The transcoding module 1770 performs the actual transcoding function including transcoding between MPEG series, H.26X video series, GSM-AMR and G.72X audio codec series. Other examples of PTS transcoding may include at least MPEG2-audio to MPEG4-audio, G.723.1 to GSM-AMR, MPEG2-video to MPEG4-video, H.263 to MPEG4-video. The rate control module 1780 performs a throttling function and prevents network segments over which the end-points communicate from being overloaded. The intellectual property rights management module 1790 protects the IP rights. For example, the management module 1790 uses data on IP rights in a bitstream, and assists on implementing mechanisms for auditing, monitoring, billing, and protection of the IP rights associated with the received and transcoded bitstream. The capabilities processing module 1792 utilizes data in received messages to find the best matching mode for the specific media that is transmitted from one end-point to another.

WO 02/073443

PCT/US02/08218

[99] The architecture of a PTS determines performance, cost, and time-to-market of the server. Performance may be considered as the number of simultaneous gateway channels or calls that the PTS may simultaneously process. Furthermore, for a fixed number of channels, the cost and performance of the architecture will depend on the following factors:

- 5        [100] 1. Bus architecture;
- [101] 2. Transcoding architecture and hardware for various video and audio transcoding;
- [102] 3. Network off loading for connecting to MGC and other gateway components; and
- 10       [103] 4. Operating System.

[104] Figure 18 shows an explanation of the symbols used in the flow chart. This diagram is merely an example, which should not unduly limit the scope of the claims herein. One of ordinary skill in the art would recognize many other variations, modifications, and alternatives.

- 15       [105] Figure 19 shows a simplified flow chart illustrating the high-level procedure for video bitstream transcoding that may be used in a PTS according to an embodiment of the present invention. This diagram is merely an example, which should not unduly limit the scope of the claims herein. One of ordinary skill in the art would recognize many other variations, modifications, and alternatives. The procedure reads a chunk of bits and if
- 20       the end-sequence marker is detected then it ends. Otherwise it reads the next code word, transcodes the code word to the output protocol codeword, updates the history record and emits the transcoded bits to the output buffer which is flushed according to a rate control scheme as to avoid that the input buffer at the receiving end-point does not over-flow.

- [106] Figure 20 illustrates a PTS hardware architecture according to an
- 25       embodiment of the present invention. This diagram is merely an example, which should not unduly limit the scope of the claims herein. One of ordinary skill in the art would recognize many other variations, modifications, and alternatives. The architecture features an intelligent transcoding node 2010, which is a bus card including one or more processors, a bank of DSP
- processors 2020, one or more network interfaces 2030, one or more processors 2040, and a
- 30       memory bank 2050. This architecture has several advantages. First, the network interface 2030 is embedded in the intelligent transcoding node bus-card 2010. Hence the call processing and

WO 02/073443

PCT/US02/08218

transcoding are performed locally on the bus-card 210. Second, one or more network interfaces per processing node are possible. Third, the architecture can support a large number of simultaneous calls because of compactness of its processing modules.

[107] In addition to the embodiment in Figure 20, numerous architectures for a  
5 PTS are possible, a few of which are listed below:

- [108] 1. Standalone chassis with bus-cards;
- [109] 2. PC-like implementation as described below;
- [110] 3. Firmware to existing processing hardware, including ASICs;
- [111] 4. Software running on existing hardware;
- 10 [112] 5. Software running on existing hardware with hardware acceleration  
by the means of ASICs, DSPs, or other types of processors; and
- [113] 6. ASIC chipset.

[114] Figure 21 shows an embodiment of a computer system according to the  
present invention. This diagram is merely an example, which should not unduly limit the scope  
15 of the claims herein. One of ordinary skill in the art would recognize many other variations,  
modifications, and alternatives. The present invention may be implemented in a personal  
computer (PC) architecture. Alternative computer system architectures, or other programmable  
or electronic-based devices may also be employed.

[115] In Figure 21 a computer system 2100 comprises a bus 2101 for  
20 communicating information, a processor 2102 coupled to the bus 2101 for processing  
information, random access memory 2103 coupled to the bus 2101 for storing information and  
instructions for the processor 2102, a read-only memory 2104 coupled to the bus 2101 for storing  
static information and instructions for the processor 2102 and PTS applications, a display device  
2105 coupled to the bus 2101 for displaying information for a user, an input device 2106 coupled  
25 to the bus 2101 for communicating information and command selections to the processor 2102,  
and a mass storage device 2107, e.g., a magnetic disk and associated disk drive, coupled to the  
bus 2101 for storing information and instructions. A data storage medium 2108 contains digital  
information, e.g., PTS software modules, and it is coupled to the bus 2101 and configured to  
operate with mass storage device 2107 to provide the processor 2102 access to digital  
30 information stored on the data storage medium 2108 through the bus 2101. A hardware  
transcoding acceleration module 2109 comprises printed circuit boards, Digital Signal

WO 02/073443

PCT/US02/08218

Processors, ASIC's, and FPGA's. The module 2109 is communicatively coupled to the bus 2101, and may have an embodiment similar to the intelligent transcoding node 2010 as shown in Figure 20.

[116] Processor 2102 may be any of a wide variety of general purpose processors or microprocessors, e.g., the Pentium™ processor manufactured by Intel Corporation, and a MIPS processor manufactured by MIPS Technologies, Inc., of 2011 N. Shoreline Blvd., Mountain View, CA 94039-7311. Other varieties of processors such as digital signal processors (DSP's) may also be used in the computer system 2100. The display device 2105 may be a liquid crystal device, cathode ray tube (CRT), or other suitable display device. The mass storage device 2107 may be a conventional hard disk drive, a floppy disk drive, a CD-ROM drive, or other magnetic or optical data storage device for reading and writing information stored on a hard disk, a floppy disk, a CD-ROM, a magnetic tape, or other magnetic or optical data storage medium. The data storage medium 2108 may be a hard disk, a floppy disk, a CD-ROM, a magnetic tape, or other magnetic or optical data storage medium.

[117] In general, the processor 2102 can retrieve processing instructions and data from the read-only memory 2104. The processor 2102 can also retrieve processing instructions and data from the data storage medium 2108 using the mass storage device 2107 and downloads the information into the random access memory 2103, which may be a SDRAM. The processor 2102 then executes an instruction stream from the random access memory 2103 or read-only memory 2104. Command selections and information input at the input device 2106 may direct the flow of instructions executed by the processor 2102. Input device 2106 may be, among others, a pointing device such as a conventional mouse or a trackball device. The execution results may be displayed on the display device 2105. The computer system 2100 also comprises a network device 2110 for connecting the computer system 2100 to a network. The network device 2110 may be an Ethernet device, a phone jack, a satellite link, or other device.

[118] Embodiments of the present invention may be represented as a software product stored on a machine-accessible medium, also referred to as a computer-accessible medium or a processor-accessible medium. The machine-accessible medium may be any type of magnetic, optical, or electrical storage medium including a diskette, a CD-ROM, a memory device, either volatile or non-volatile, an ASIC, a firmware to an ASIC, a system-on-chip, or other storage mechanism. The machine-accessible medium may contain various sets of

WO 02/073443

PCT/US02/08218

instructions, code sequences, or configuration information. Other data necessary to implement the present invention may also be stored on the machine-accessible medium. As merely an example, transcoding techniques are described in U.S. Provisional Serial No. 60/347270 (Attorney Docket No. 021318-000200US), commonly assigned, and hereby incorporated by reference for all purposes.

[119] In addition to media transcoding that the PTS can perform, it can also perform system protocol transcoding. A multimedia system protocol is typically an umbrella of protocols that define how multimedia end-points can connect to each other, can issue and interpret commands (such streaming or opening a video channel), can tear down connections, can join conferences. A system protocol typically covers the following important aspects: call signalling, command and control, media transports aspects, and media coding aspects. For example the H.323 system protocol standard covers H.225.0/Q.931 for call signalling and media transports, H.245 for command and control, and a number of audio and video codecs. The H.324 system protocol standard covers H.223 (media and data bitstreams multiplexing), H.245 for command and control, and a number of audio and video codecs. Although some aspects of H.323 and H.324 are similar, H.323 is packet based whereas H.324 is circuit based. Hence for H.323 and H.324 end-points to communicate, system protocol transcoding needs to be performed, at the call-signaling level, command and control levels and media coding levels. The PTS performs both system protocol transcoding and media (audio and video) transcoding. From the PTS point of view, the Call signaling transcoding is the process of proxying the end-points (H.324, H.323, SIP, RTSP, and others) so they can connect to each other. From the command and control point of view, the PTS performs the transcoding so messages such as terminal capabilities, open/close logical channels, etc., are translated so they can be understood by the terminal receiving the command and control message. The PTS has to perform translation of the issued messages it receives from the sending end-point so they could be understood by the receiving terminal. In terms of media transport, the PTS needs to access data by demultiplexing from circuit bearer channels, extracting the media service data units, transcoding the media bits and then packaging the transcoded bits so they can be sent in a format the receiving end point can understand. If the receiving end point is H.323, the packaging would involve RTP packetization. So the media transport transcoding (translation) consists of circuit-to-packet, packet-to-circuit and packet-to-packet translation of the media bits. When the transport is circuit

WO 02/073443

PCT/US02/08218

based, the typically the media bits are multiplexed in a Time Domain Multiplexing (TDM) fashion.

{120} While there has been illustrated and described what are presently considered to be example embodiments of the present invention, it will be understood by those skilled in the art that various other modifications may be made, and equivalents may be substituted, without departing from the true scope of the invention. Additionally, many modifications may be made to adapt a particular situation to the teachings of the present invention without departing from the central inventive concept described herein.

WO 02/073443

PCT/US02/08218

**WHAT IS CLAIMED IS:**

- 1           1.     A system for transferring multimedia information from a source location  
2 to a destination location through one or more networks, the system comprising:  
3           a source output providing a first stream of information in one of a plurality of  
4 source;  
5           a destination input receiving a second stream of information in one of a plurality  
6 of destination capabilities;  
7           a proxy transcoder server ("PTS") coupled between the source output and the  
8 destination input, the PTS comprising:  
9           a capability module adapted to identify the source capability of the source  
10 output and adapted to identify destination capability of the destination input;  
11           a selection module adapted to select a transcoding process based upon the  
12 one capability of the source capabilities and the one capability from the destination  
13 capabilities;  
14           a transcoding module adapted to use the selected transcoding process to  
15 process the first stream of information.
- 1           2.     The system of claim 1 wherein the one or more transport networks are  
2 selected from a group comprising the Internet, a mobile network, a wide area network, a local area  
3 network, PTSN, ISDN, and SONET.
- 1           3.     The system of claim 1 wherein at least one of the source output and the  
2 destination input is that of a remote device.
- 1           4.     The system of claim 3 wherein the capability module identifies at least one  
2 of the output and input of the remote device, based on information stored in the device, based on  
3 user subscription information stored in a network database of the user's service provider, based  
4 on in-band information command and control within a stream exchanged, or pre-set by the  
5 service provider.



WO 02/073443

PCT/US02/08218

- 1           5.     The system of claim 1 wherein the transcoding process selected by the  
2     capability module transcodes data from a first bitstream protocol mode to a second bitstream  
3     protocol mode.
- 1           6.     The system for claim 1 wherein the PTS further comprising a rate control  
2     module regulating the data rate produced by the PTS.
- 1           7.     The system for claim 6 wherein the rate control module detects network  
2     status information by calculating "round-trip" time information based on network congestion  
3     information, bandwidth information, quality information from a network host or network access  
4     provider, or internal PTS mechanisms.
- 1           8.     The system for claim 7 wherein the "round-trip" time information can be  
2     measured by send a "ping" packet to either the source location or the destination location.
- 1           9.     The system for claim 6 wherein the rate control module detects the  
2     network status information by using in-band information.
- 1           10.    The system for claim 6 wherein the rate control module regulates the data  
2     rate by changing transcoding parameters.
- 1           11.    The system for claim 6 wherein the rate control module regulates the data  
2     rate by instructing network equipment to give a higher priority to data being handled by the PTS  
3     than other data.
- 1           12.    The system of claim 1 wherein the format of the capability is selected  
2     from a group comprising ITU, IETE, and WAP.
- 1           13.    The system of claim 1 wherein the one or more networks are selected from  
2     a plurality of different networks, each of the network being configured for a particular standard.
- 1           14.    The system of claim 1 wherein the PTS further comprising a network  
2     addressing module to determine the network address of the source output and the network  
3     address of the destination input.

WO 02/073443

PCT/US02/08218

1           15. The system of claim 1 wherein the PTS further comprising a media mixing  
2 process to combine bitstreams associated with two or more audio streams and retransmit the  
3 combined bitstreams to the destination input.

1           16. The system of claim 1 wherein the PTS further comprising an intellectual  
2 property rights management module to manage and process information on intellectual property  
3 rights.

1           17. The system of claim 1 wherein the PTS further comprising a encryption  
2 and decryption process to encrypt and decrypt the data.

1           18. The system for claim 6 wherein the rate control module regulates the data  
2 rate dynamically and in real time.

1           19. The system of claim 1 wherein the transcoding module are programmable  
2 to transcode between various types of capabilities for the source output and various types of  
3 capabilities for the destination input.

1           20. A system for transferring multimedia information from source to  
2 destination locations through one or more networks, the system comprising:  
3           a source output in a first format from a plurality of source capabilities, the source  
4 output being coupled to a first network, the source output providing a first stream of information;  
5           a destination input to be received in a second format from a plurality of  
6 destination capabilities, the destination input being coupled to a second network, the destination  
7 input receiving a second stream of information;  
8           a proxy transcoder server ("PTS") coupled between the source output and the  
9 destination input, the proxy transcoder server comprising:  
10           a capability process coupled to the source output, the capability process  
11 being adapted to identify the first format of the source output and adapted to identify the second  
12 format of the destination input;  
13           a transcoding process coupled to the capability process, the transcoding  
14 process comprising a plurality of transcoding modules numbered 1 through N, where N is an  
15 integer greater than 1, the transcoding process being adapted to select one of the transcoding

WO 02/073443

PCT/US92/08218

16 process based upon the first format that is associated with a capability and the second format that  
17 is associated with a second capability; and

18 a bit rate control process coupled to the transcoding process, the bit rate  
19 control process being adapted to receive a network status information from the first network, the  
20 bit rate control being adapted to adjust a status of the stream of information based upon the  
21 network status information.

1 21. The system of claim 20 wherein the status information comprises a ping.

1 22. The system of claim 20 wherein the status is a stop status.

1 23. The system of claim 20 wherein the status is a prioritization status.

1 24. The system of claim 20 wherein the status is to adjust a bit rate by  
2 selecting a lower bit rate coder.

1 25. A method for processing streams of information, the method comprising:  
2 identifying a source capability from a plurality of source capabilities for a stream  
3 of information;

4 identifying a destination capability from a plurality of destination capabilities;

5 selecting a transcoding process from a plurality of transcoding processes in a

6 library based upon the identified source capability and the identified destination capability;

7 processing the stream of information using the selected transcoding process if the  
8 identified source capability and the identified destination capability are different;

9 transferring the stream of information from the source to the destination free from  
10 one of the transcoding processes of the identified source capability and the identified destination  
11 capability matches.

1 26. The method of claim 25 wherein the selected transcoding process is  
2 provided by empirical information.

1 27. The method of claim 25 wherein the library is a look up table having at  
2 least the plurality of source capabilities and the plurality of destination capabilities in a second  
3 dimension.

WO 02/073443

PCT/US02/08218

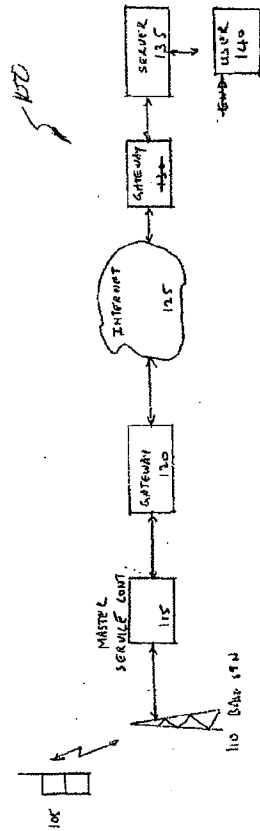


Figure 1 (PrioA Net)

WO 02/073443

PCT/US02/08218

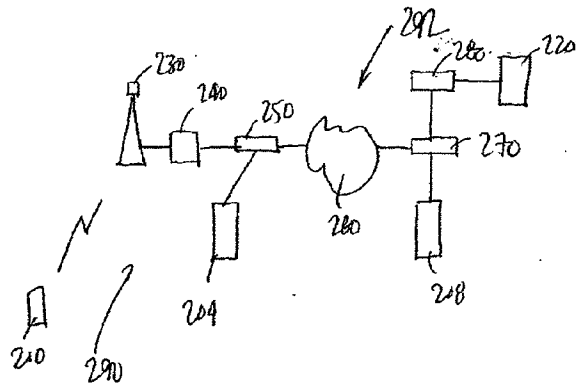


FIG. 2

WO 02/073443

PCT/US02/08218

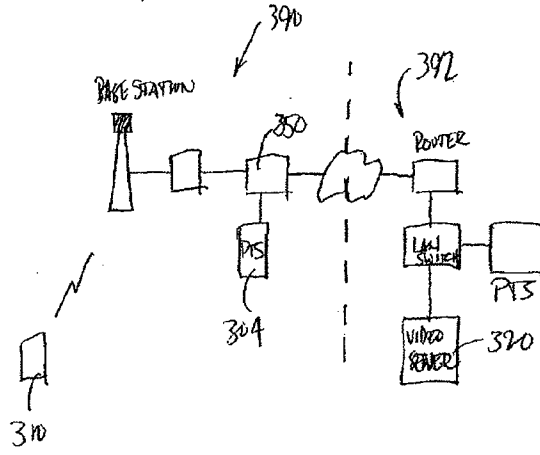


FIG. 3

WO 02/073443

PCT/US02/08218

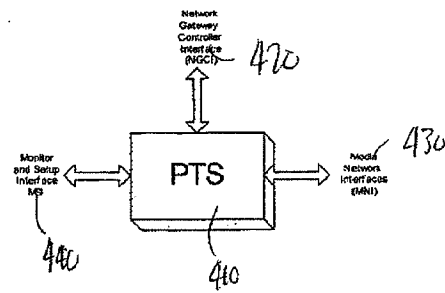


FIG. 4

WO 02/073443

PCT/US02/08218

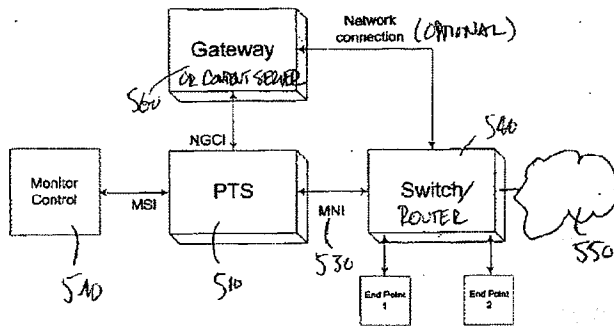


FIGURE 5



WO 02/073443

PCT/US02/08218

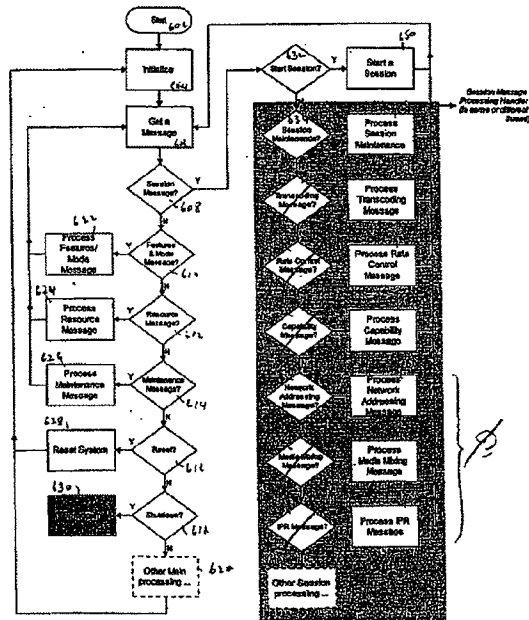


FIG. 6

6/15

WO 02/073443

PCT/US92/08218

## Main System Messages (A)

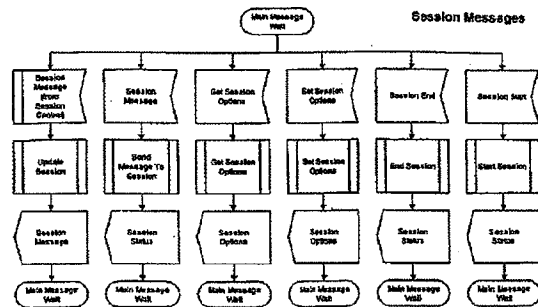


FIGURE 7.

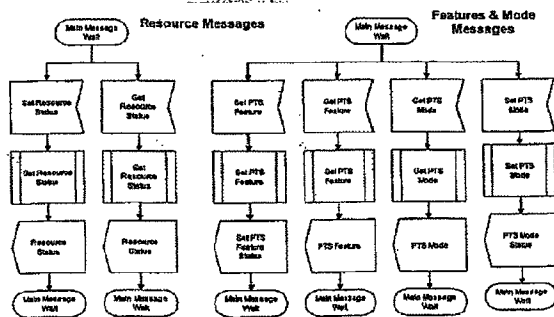


FIGURE 9.

FIGURE 8.

WO 02/073443

PCT/US02/08218

Figure 6. Main procedure messages.

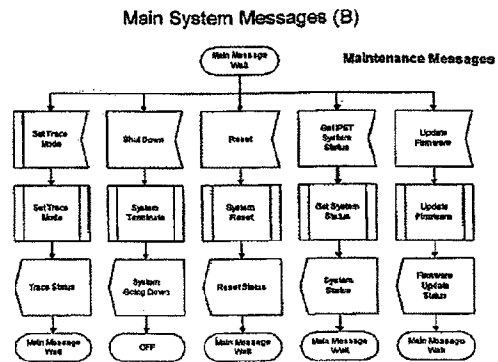


FIGURE 10

## Figure 2.1

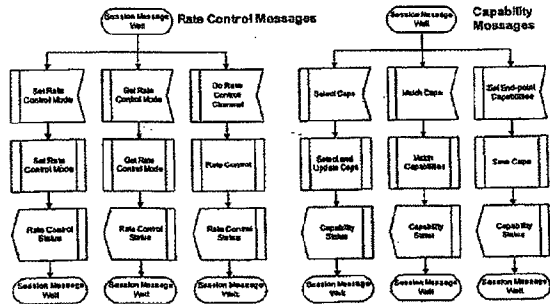


Figure 13

WO 02/073443

PCT/US02/08218

## Session Messages (B)

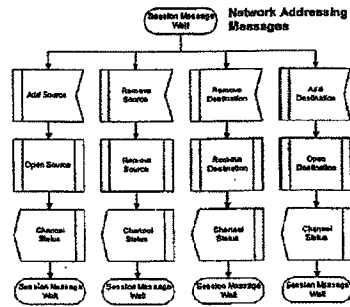


FIGURE 14

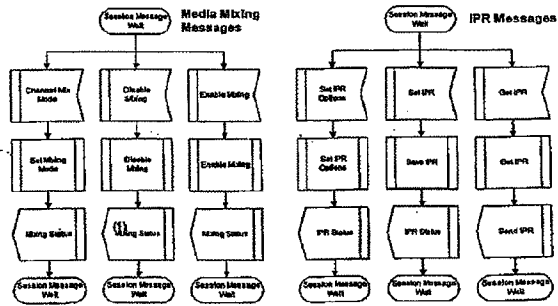


FIGURE 15

FIGURE 16

WO 02/073443

PCT/US02/08218

## PTS Software Modules

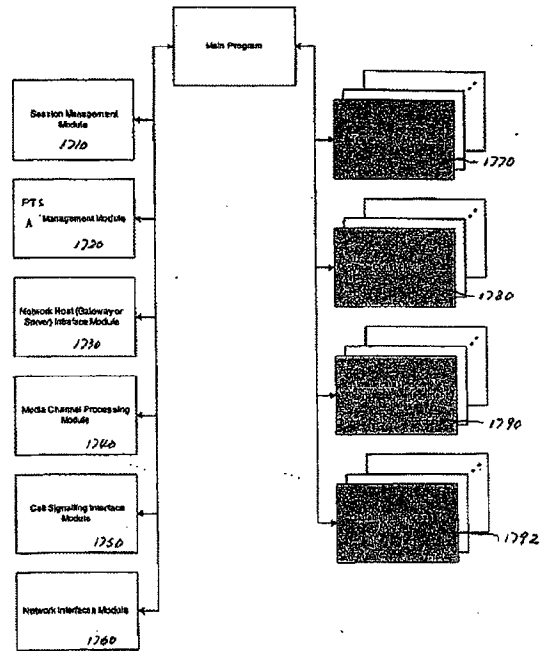
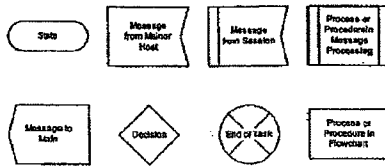


Figure 17

WO 02/073443

PCT/US02/08218

## Symbols



Symbols used in flowcharts and message diagrams.

FIGURE 18

WO 02/073443

PCT/US92/08218

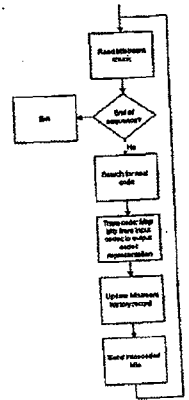


Figure 19



WO 02/073443

PCT/US02/09218

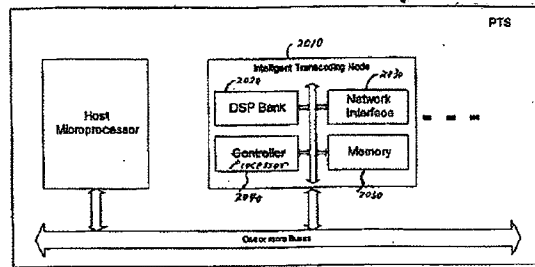


FIGURE 20

WO 02/073443

PCT/US02/08218

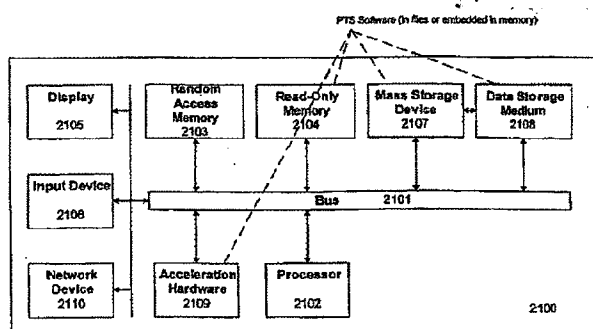


FIGURE 21

## 【国際公開パンフレット (コレクトバージョン)】

(32) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

CORRECTED VERSION

(19) World Intellectual Property  
Organization  
International Bureau(43) International Publication Date  
19 September 2002 (19.09.2002)

PCT

(10) International Publication Number  
WO 2002/073443 A1

(51) International Patent Classification: G06F 15/16, H03M 3/00

(74) Agents: OGAWA, Richard, T. et al.; Townsend and Townsend and Crew LLP, 2 Embarcadero Center, 8th Floor, San Francisco, CA 94111 (US).

(21) International Application Number: PCT/US2002/008218

(22) International Filing Date: 13 March 2002 (13.03.2002)

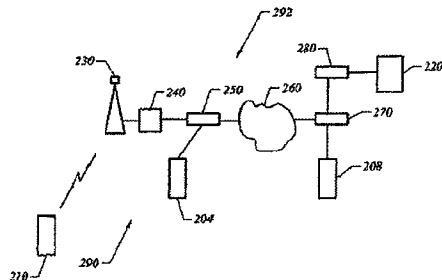
(23) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
60215,584 13 March 2001 (13.03.2001) US  
10/599,901 12 March 2002 (12.03.2002) US(71) Applicant: MACCERNA PTY LTD. [AU/AU]; Level 7,  
Stratford Street, Broadway, NSW 2007 (AU).(72) Inventor: JABRI, Macwan, Anwar; Unit 44, 367 Que-  
ensway Street, Sydney, New South Wales 2000 (AU).(81) Designated States (national): AE, AG, AL, AM, AT, AU,  
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,  
CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GR, GU, HK,  
HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,  
LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW,  
MX, MY, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG,  
SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU,  
ZA, ZM, ZW.(84) Designated States (regional): ARIPO patent (GH, GM,  
KE, LS, MW, MZ, SD, SI, SZ, TZ, UG, ZM, ZW),  
Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),  
European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR,  
GB, GR, IE, IT, LI, MC, NL, PT, SE, TR), OAPI patent  
(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR,  
NE, SN, TD, TG).

(Continued on next page)

(54) Title: METHOD AND SYSTEM FOR TRANSCODING VIDEO AND SPEECH SIGNALS



(57) Abstract: A system and method for transcoding multimedia information from a source location (210) to a destination location (260) through one or more networks (230, 270), which may be different. The system has a source output which provides a first stream of information in a first format. The system also has a destination input which receives a second stream of information in a second format. A proxy transcoding server ("PTS") (204, 208) is coupled between the source output and the destination input. The PTS has a transcoding module transcoding data. The PTS has a transcoding module transcoding data. The PTS also has a capability module identifying a first capability of the source output and a second capability of the destination input, and selecting a transcoding process based upon the first capability and the second capability.

WO 2002/073443 A1

WO 2002/073443 A1 

Published:  
— with international search report

(15) Information about Correction:  
see PCT Gazette No. 15/0004 of 8 April 2004, Section II.  
For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(48) Date of publication of this corrected version:  
8 April 2004

WO 2002/073443

PCT/US2002/008218

**METHOD AND SYSTEM FOR TRANSCODING VIDEO AND SPEECH  
SIGNALS**

5 **CROSS-REFERENCES TO RELATED APPLICATIONS**

[01] The present application claims priority to U.S. Provisional Patent Application No. 60/275,584 filed March 13, 2001, which is commonly assigned, and hereby incorporated by reference for all purposes.

10 **STATEMENT AS TO RIGHTS TO INVENTIONS MADE UNDER  
FEDERALLY SPONSORED RESEARCH OR DEVELOPMENT**

[02] NOT APPLICABLE

15 **REFERENCE TO A "SEQUENCE LISTING," A TABLE, OR A COMPUTER  
PROGRAM LISTING APPENDIX SUBMITTED ON A COMPACT DISK.**

[03] NOT APPLICABLE

**BACKGROUND OF THE INVENTION**

20 [04] The present invention is related to the field of communications. In particular, the present invention is related to a method and apparatus for transcoding video and audio signals. More particularly, the invention provides a method and system for transcoding information (e.g., video, voice, data) from a first format to a destination format using a proxy transcoder server having a plurality of transcoding processes, where at least one is selected for transcoding the information. Merely by way of example, the invention is applied to a wide area telecommunication network, but it would be recognized that the invention can also be applied across many different types of multimedia protocols over transport networks such as the Internet, a mobile network, a local area network, PTSN, ISDN, SONET, DWDM, and others.

25 [05] Telecommunication techniques have improved dramatically over the years. Many different types of networks such as fixed switched, packet based, wireless and

WO 2002/073443

PCT/US2002/006218

mobile have been deployed. One of the most widely known world wide network called the "Internet" has popularized networking to many people around the world. As increase in use of wide area networks such as the Internet has resulted in many new on-line services such as electronic mail, video telephony, video streaming, electronic commerce, and others. Although computers originally connected to the Internet, other devices such as mobile phones, personal digital assistants, laptop computers, and the like have also been connected. Accordingly, many different types of devices now have access to many different types of services over a variety of networks.

[06] A variety of network elements make up the networks, which connect the aforementioned devices together. Such devices are often connected by gateways and switches that handle transfer of data and conversion of messages from protocols of a sending network to protocols used by a receiving network. Gateways and switches convert analog voice messages to digital formats including G.711 and G.723.1, which are ITU standards. Gateways transmit the converted messages typically in a way similar to transmission of voice over IP. G.711 is an ITU standard for speech codecs that provides audio signals at 64 Kbps using either the A-Law PCM method or the mu-Law PCM method. G.723.1 is an ITU standard for speech codecs optimized for narrow-band networks, including Plain Old Telephone Systems and narrow band Internet connections. The standard uses the LD-CELP method and provides audio signals at 5.3 or 6.3 Kbps. Depending upon the application, there can be many others as well.

[07] As merely an example in Figure 1, a conventional system 100 is shown. This diagram is merely an example and provided for illustrative purposes only. A message originates from a mobile device 105, which is coupled to a wireless network. The message is sent from the mobile device to base station 110 through the wireless network. The base station is coupled to a service station 115, which is coupled to gateway 120. The base station receives the radio message from the mobile device 105, and converts the message, without transcoding, into a digital format, and transmit it to the service station 115. The reformatted message is subsequently transmitted to the gateway which in turn transmits the message to its destination, a user 140, through the Internet 125 and also through a variety of network elements. Such elements may include a gateway 130, a server 135, and others.

[08] One or more gateways may also convert videoconferencing signals from one digital format to another, such as from H.320 to H.323, and transmit converted signals over

WO 2002/073443

PCT/US2002/008218

the Internet. H.320 is an ITU standard for videoconferencing over digital lines, and it uses the H.261 video compression method, which allows H.320-compliant videoconferencing and desktop systems to communicate with each other over ISDN, switched digital lines and leased lines. H.323 is an ITU standard for real-time, interactive voice and videoconferencing over LANs and the Internet. Widely used for IP telephony, H.323 allows any combination of voice, video and data to be transported. H.323 specifies several video codecs, including H.261 and H.263, and audio codecs, including G.711 and G.723.1. Unfortunately, the audio and video standards have grown well beyond H.320, H.323, G.711, and G.723.1. That is, the proliferation of different standards has caused difficulty in communicating messages between them.

Additionally, any communication between such standards has caused a proliferation of complex conversion techniques, which are time consuming and lack efficiency. Accordingly, there is a need for an efficient way to convert information or transcode between various formats in real time. Because some systems such as H.320 and H.324 are circuit switched systems (data is transmitted as a continuous stream of bits) and some other systems are packet based, the connection of circuit-based to packet based systems require the demultiplexing of bits from circuit based bitstreams into packet (circuit-to-packet) and vice versa (packet-to-circuit). Note that different system protocols such as H.320, H.323, H.324, 3GPP-324M, SIP and SDP, make use of different signaling methods (to setup connections and exchange terminal capabilities). The interconnectivity of these systems require the trans-signaling and the conversion of terminal capabilities so terminal can understand what terminals using different protocols are capable of.

{09} From the above, it is seen that an improved way of transferring information from a source to a destination is highly desirable.

WO 2002/073443

PCT/US2002/066218

## SUMMARY OF THE INVENTION

[10] According to the present invention, improved techniques for transcoding in the telecommunication fields are provided. In particular, the present invention is related to a method and apparatus for transcoding video and speech signals. More particularly, the invention provides a method and system for transcoding information (e.g., video, voice, data) from a first format to a destination format using a proxy transcoder server having a plurality of transcoding processes, where at least one is selected for transcoding the information. Merely by way of example, the invention is applied to a wide area telecommunication network, but it would be recognized that the invention can also be applied across many different types of multimedia protocols over transport networks such as the Internet, a mobile network, a local area network, PTSN, ISDN, SONET, DWDM, and others.

[11] In a specific embodiment, the invention provides a system for transferring multimedia information from a source location to a destination location through one or more networks, which may be different. The system has a source output which provides a first stream of information in a first format. The system also has a destination input which receives a second stream of information in a second format. A proxy transcoder server ("PTS") is coupled between the source output and the destination input. The PTS has a transcoding module transcoding data. The PTS also has a capability module identifying a first capability of the source output and a second capability of the destination input, and selecting a transcoding process based upon the first capability and the second capability. Preferably, the selecting is provided using capability mode selection.

[12] In an alternative specific embodiment, the invention provides a system for transferring multimedia information from source to destination locations through one or more networks, which may be different. The system has a source output in a first format, where the source output is coupled to a first network, the source output providing a first stream of information. The system also has a destination input to be received in a second format, where the destination input is coupled to a second network. The destination input receives a second stream of information. A proxy transcoder server ("PTS") is coupled between the source output and the destination input. The proxy transcoder server has a capability process, which is adapted to identify a first capability of the source terminal (which may have different capabilities) and is



WO 2002/07343

PCT/US2002/006218

adapted to identify a second capability of the destination terminal (which may also have different capabilities). The server also has a transcoding process comprising a plurality of transcoding modules numbered 1 through N, where N is an integer greater than 1. The transcoding process is adapted to selected one of the transcoding process based upon the first capability and the second capability. The proxy transcoder server has a bit rate control process. The bit rate control process is adapted to receive a network status information (e.g., ping) from the first network. The bit rate control is adapted to adjust a status (e.g., stop, prioritize allow, adjust bit rate (by selecting lower bit rate coder)) of the stream of information based upon the network status information.

[13] In an alternative specific embodiment, the invention provides a method for processing streams of information. The method includes identifying a source capability from a plurality of source capabilities for a stream of information. The method also identifies a destination capability from a plurality of destination capabilities. A step of selecting a transcoding process from a plurality of transcoding processes in a library based upon the identified source capability and the identified destination capability is included. The method also processes the stream of information using the selected transcoding process if the identified source capability and the identified destination capability are different. The method also transferring the stream of information from the source to the destination free from one of the transcoding processes if the identified source capability and the identified destination capability matches.

[14] Numerous benefits are achieved using the present invention over conventional techniques. In a specific embodiment, the invention provides a way to transcode video data from H.263 to MPEG-4 data and vice versa (and between other video codecs), or to transcode audio data from G.723.1 to GSM-AMR (and between other audio codecs) and vice versa. Transcoding can proceed seamlessly so that end points receiving transcoded data do not notice the conversion in preferred embodiments. The invention can also be implemented using conventional software and hardware technologies, such as digital signal processors (DSPs). Depending upon the embodiment, one or more of these benefits or features can be achieved. These and other benefits are described throughout the present specification and more particularly below.

WO 2002/073443

PCT/US2002/049218

[15] The accompanying drawings, which are incorporated in and form part of the specification, illustrate embodiments of the invention and, together with the description, serves to explain the principles of the invention.

## BRIEF DESCRIPTION OF THE DRAWINGS

- 5 [16] Figure 1 illustrates a simplified block diagram of a mobile phone communicating with an end-user.
- [17] Figure 2 illustrates one embodiment of the operation of proxy transcoder servers:
- 10 [18] Figure 3 illustrates another embodiment of the operation of proxy transcoder servers.
- [19] Figure 4 is a simplified block diagram illustrating one embodiment of the connectivity of the proxy transcoder server.
- [20] Figure 5 is a simplified block diagram illustrating an embodiment of a proxy transcoder server connected to a gateway.
- 15 [21] Figure 6 is a simplified flow diagram illustrating one embodiment of the transcoding process.
- [22] Figure 7 is a simplified flow diagram illustrating one embodiment of the main system messages.
- [23] Figure 8 is a simplified flow diagram illustrating one embodiment of the 20 resource messages.
- [24] Figure 9 is a simplified flow diagram illustrating one embodiment of the PTS features and mode messages.
- [25] Figure 10 is a simplified flow diagram illustrating one embodiment of the PTS maintenance messages.
- 25 [26] Figure 11 is a simplified flow diagram illustrating one embodiment of the PTS session maintenance and transcoding messages.
- [27] Figure 12 is a simplified flow diagram illustrating one embodiment of the PTS session rate control messages.
- [28] Figure 13 is a simplified flow diagram illustrating one embodiment of the 30 PTS session capability messages.

WQ 2002/073443

PCT/US2002/008218

[29] Figure 14 is a simplified flow diagram illustrating one embodiment of the network addressing messages.

[30] Figure 15 is a simplified flow diagram illustrating one embodiment of the PTS media mixing messages.

5 [31] Figure 16 is a simplified flow diagram illustrating one embodiment of the PTS JPR messages.

[32] Figure 17 is a simplified flow diagram illustrating one embodiment of the PTS software modules.

10 [33] Figure 18 is a simplified flow diagram illustrating symbols used in the flow diagrams.

[34] Figure 19 is a simplified flow diagram illustrating one embodiment of the transcoding procedure.

[35] Figure 20 is a simplified block diagram illustrating one embodiment of a PTS's hardware architecture.

15 [36] Figure 21 is a simplified block diagram of a computer system that may be used to implement an embodiment of the invention.

#### DESCRIPTION OF THE SPECIFIC EMBODIMENTS

20 [37] According to the present invention, improved techniques for transcoding in the telecommunication fields are provided. In particular, the present invention is related to a method and apparatus for transcoding video and audio signals. More particularly, the invention provides a method and system for transcoding information (e.g., video, audio, data) from a first format to a destination format using a proxy transcoder server having a plurality of transcoding processes, where at least one is selected for transcoding the information. Merely by way of  
25 example, the invention is applied to a wide area telecommunication network, but it would be recognized that the invention can also be applied across many different types of multimedia protocols over transport networks such as the Internet, a mobile network, a local area network, PTSN, ISDN, SONET, DWDM, and others.

30 [38] Parts of the description will be presented in terms of operations performed through the execution of programming instructions according to various embodiments. As will be understood by those skilled in the art, these operations often take the form of electrical,

WO 2002/073443

PCT/US2002/008218

magnetic, or optical signals capable of being stored, transferred, combined, and otherwise manipulated through, for instance, electrical components. Parts of the description will be presented using a distributed computing environment. In a distributed computing environment, file servers, computer servers, and memory storage devices may be located in different places, but they are accessible to local processing units through the network. In addition, program modules may be physically located in different local and remote memory storage devices. Execution of the program modules may occur locally in a stand-alone manner or remotely in a client-server manner. Examples of such distributed computing environments include local area networks of an office, enterprise-wide computer networks, and the global Internet.

[39] Additionally, the following terms are provided to assist the reader in explaining aspects of the invention. Such terms are not intended to be limiting but merely provided for descriptive purposes to one of ordinary skill in the art. Other meanings for the terms consistent with those understood by one of ordinary skill in the art may also be used.

Term	Description
ASIC	Application Specific Integrated Circuit
CIF	Common Intermediate Format
ETSI	European Telecommunications Standards Institute
G.723.1	ITU Recommendation G.723.1, Dual rate speech coder for multimedia communications transmitting at 5.3 and 6.3 kbit/s, 1996
GOB	Group of blocks
GSM	Global System for Mobile communications
GSM-AMR	ETSI Adaptive Multi Rate speech coder. GSM 06.90: "Digital cellular telecommunications system (Phase 2+); AMR speech transcoding", 1998
GSM-AMR	GSM- Adaptive Multi Rate
H.320	ITU Recommendation H.320, Narrow-band visual telephone systems and terminal equipment, 1997

WO 2002/073443

PCT/US2002/009218

H.323	ITU Recommendation H.323, Packet-based multimedia communications systems, 1998
H.324	ITU Recommendation H.323, Terminal for low bit-rate multimedia communication, 1998
H.261	ITU Recommendation H.263, Video codec for audiovisual services at $p \times 64$ kbit/s, 1993
H.263	ITU Recommendation H.263, Video coding for low bit rate communication, 1998
IETF	Internet Engineering Task Force
ISO	International Standard Organisation
ITU	International Telecommunication Union
MB	Macro block
MPEG	Moving Picture Expert Group, part of the International Standard Organisation
MPEG2	MPEG audiovisual standards 13818 series
MPEG4	MPEG audiovisual standards 14496 (1-5)
MVD	Motion Vector Data
P Frame or P Picture	Video frame based on predicted information
PTS	Proxy Transcoder Server
QCIF	Quarter CIF (see CIF)
RFC	Request for Comment
SDP	Session Description Protocol
SIP	Session Initiation Protocol
TCOEF or TCOEFF	Transform Coefficients
W3C	World Wide Web Consortium
WAP	Wireless Access Protocol

[40] In a specific embodiment, the term "Proxy Transcoder Server" (herein PTS) is a computer having various modules to carryout some or all of the functionality described herein as well as other functionality, which are known or not known at this present time. The

WO 2002/073443

PCT/US2002/068218

PTS comprises a host processor, one or more network interfaces, and one or more transcoders.

A transcoder may contain, among others, a printed circuit board, an application specific integrated circuit (ASIC), and a field programmable gate array (FPGA). The PTS may be connected to a network host such as a media gateway controller or a soft switch or to a content

5 server such as a video server. The PTS can provide codecs and interfaces to various network architectures and protocols, including WAN, LAN, Mobile, PDSN, ISDN, SONET, and perform one or more of the following functions:

- [41] 1. Capability matching and mode selection;
- [42] 2. Media bitstream transcoding;
- 10 [43] 3. Media bitstream rate control;
- [44] 4. Intellectual property rights management and processing;
- [45] 5. Audio data mixing; and
- [46] 6. Encryption and/or decryption
- [47] The overall function of the PTS is to translate between various protocols,
- 15 as exemplified below.

[48] 1. Translating multimedia systems protocols, including the ITU H.32X series, e.g., H.242 and H.245, which are used for conveying the media capabilities of end-points, and managing media channels and conferences.

[49] 2. Translating media streams including audio and video streams, e.g.,  
20 translating video streams to and from any pair of video codecs including MPEG2, MPEG4, H.261, H.263, or audio streams to and from any pair of audio codecs including G.723.1, G.729, GSM-AMR, EVRC, SMV, and QCELP.

[50] 3. Translating IP management protocols and IP rights markings. For example, the PTS may extract IP rights information from MPEG4 streams and process the  
25 information in accordance with IP rights-related operations in order to maintain the IP rights in the media.

[51] 4. Encrypting and/or decrypting signals if necessary.

[52] Hence, the PTS has a variety of audio and video transcoding capabilities.

Selection of capabilities affects overall qualities of service (QoS) in multimedia communication,  
30 and should therefore be determined based on the connection bandwidth and load of the gateway to which the PTS is coupled. The PTS can transcode for different media contents including

WO 2002/073443

PCT/US2002/008218

MPEG series, H.26X video series, GSM-AMR, and G.72X audio codec series. Furthermore, the PTS' transcoding capabilities can be easily upgraded, at least in part due to its programmability.

5 [53] In addition, the PTS can perform media bitstream rate control. The rate control is necessary because the connection between two end-points may involve bandwidth reduction, e.g., from wired to wireless network. The PTS may perform rate control using information from network channel allocation as supplied by the network access provider, or using in-band bandwidth management requests, which can be originated from end-points through command and control protocols, e.g., H.242 and H.245.

10 [54] Moreover, the PTS can perform intellectual property (IP) rights management and processing. For example, the PTS may identify a data set regarding intellectual property rights, and use it to facilitate the management and processing. In one embodiment, MPEG4 supplements coded media objects with an optional intellectual property identification (IP) data set, which carries information on contents, types of contents, and IP rights holders.

15 The data set, if present, forms part of an elementary stream descriptor describing streaming data associated with a media object. The number of data sets associated with each media object may vary, and different media objects may share the same data set. The provision of the data sets allows the implementation of mechanisms for audit trail, monitoring, billing, and copy protection.

20 [55] Multimedia communication applications have wide ranges of requirements for IP rights protection and security. Some applications require protecting information that users exchange in order to preserve privacy, even if the information has no intrinsic value. Other applications require high-grade management and protection for information of importance to its creator and/or distributors. In addition, the framework of IP rights management and processing must generally be flexible so as to access various forms of IP rights data, as required by a specific application, and as stored in a particular bitstream.

25 [56] Although the description that follows shows a PTS connected to a gateway as a stand alone device, alternate embodiments of the PTS may use a PTS connected to a switch, a server, a router or any device connected to a network. In addition, a PTS may be integrated into a gateway, a switch, a router or other device connected to a network to form an integral part of that network device. Further details of other embodiments are provided below.

30

WO 2002/073443

PCT/US2002/008238

[57] Figure 2 illustrates an embodiment wherein a PTS transcodes signals between a mobile end-point and a LAN phone. This diagram is merely an example, which should not unduly limit the scope of the claims herein. One of ordinary skill in the art would recognize many other variations, modifications, and alternatives. A mobile end-point 210 communicates audio and video information with a LAN phone 220. The transmitted information travels from the mobile end-point 210 to a Base Station 230 via a radio communication channel, e.g., an air link, and subsequently to a Master Service Controller 240, to a near-end gateway 250, through the Internet 260, to a far-end gateway 270, to a LAN switch 280, and finally to the LAN phone 220.

[58] Because of limitations of an air-link and mobility of the mobile end-point 210, the bandwidth between the mobile end-point 210 and the base station 230 may reach several tens or hundreds of Kbps, much smaller than the bandwidth between Internet routers and the LAN phone 220. The latter bandwidth reaches 10s to 100s of megabits per second (Mbps). In Figure 2, the media coding and decoding capability of the mobile end-point 220 is GSM-AMR for audio signals and MPEG-4 for video signals. In contrast, the media coding and decoding capability of the LAN phone 230 is G.723.1 for audio and H.263 for video. The differences in capability and bandwidth make important transcoding performed by PTS in the path between the mobile end-point 220 and the LAN phone 230, to facilitate substantially real-time communications. In addition, transcoding by PTS prevents overloading a mobile network 290, because the PTS performs a throttling function.

[59] In Figure 2, audio signals need to be transcoded between G.723.1 and GSM-AMR, while video signals need to be transcoded between MPEG4-video and H.263. Such transcoding may be performed by either a PTS 204 in the mobile network 290 or a PTS 208 in a landline network 292. Thus the two PTS's 204 and 208 are not necessary to perform the transcoding. But implementing one PTS in each network allows the bandwidth requirement at each gateway to become smaller. In addition, even though the embodiment illustrated in Figure 2 shows transcoding by PTS, the PTS's can also allow transmission of signals without transcoding.

[60] Figure 3 illustrates a block diagram of yet another embodiment of the present invention. This diagram is merely an example, which should not unduly limit the scope of the claims herein. One of ordinary skill in the art would recognize many other variations,



WO 2002/073443

PCT/US2002/008218

modifications, and alternatives. A mobile handset 310 communicates information with a video content server 320. The video server 320 streams movies containing audio and video signals. The audio signals are encoded using MPEG2-Audio Level 3 (MP3) and the video signals are encoded using MPEG2-video. Due to the mismatch in bandwidth and capability between the video server 320 and the mobile handset 310, the PST 304 in a mobile network 390 would transcode the audio to GSM-AMR and the video to MPEG4-video. But if the PST 308 in a video server network 392 performs the transcoding, the bandwidth requirement for a gateway 350 in the mobile handset network 390 may be reduced.

[61] In yet another embodiment, the capabilities of two end-points of communication may be specified using Wireless Access Protocol (WAP) User Agent Profile information (Capabilities and Preferences Information), the web consortium CC/PP (using the Resource Description Framework, or RDF), the IETF standards (RFC 2506, RFC 2533, and RFC 2703) or ITU's H.245 or H.242 standards, or a combination thereof. In yet another embodiment, a gateway may, under the assistance of PTS, detect the capabilities of end-points and the bandwidth available or allocated for transmission of media. The PTS may then select a data coding mode and transcode the media to best meet the requirements of the end-points. Hence, the PTS may serve as a proxy for each end-point side, converting bitstreams in one form appropriate for one side to another form appropriate for the other side.

[62] Figure 4 illustrates a block diagram of an PTS according to an embodiment of the present invention. This diagram is merely an example, which should not unduly limit the scope of the claims herein. One of ordinary skill in the art would recognize many other variations, modifications, and alternatives. Logical ports are defined as abstract communication ports. Several logical ports may exist on one physical network interface, but a PTS may have more than one physical network interface. A network gateway controller interface (NGCI) 420 may comprise at least one NGCI logical port for messaging between a PTS 410 and a gateway or a network host. The gateway may perform call signaling between end-points. The gateway may be a media gateway controller, or any other gateway device that establishes call or transport function between end-points. Call signaling is the process of establishing the initial link between end-point entities, by exchanging network addresses of the source and the destination of a media stream to be transcoded. For real-time Internet protocol, call signaling establishes Internet protocol addresses and Internet protocol port numbers for the

WO 2002/073443

PCT/US2002/008218

source and the destination of a media stream. Call signaling could also involve more sophisticated processes, such as that required by SIP or H.323.

[63] Media network interfaces (MNI) 430 provide logical ports for reception and transmission of media bitstreams. Through these logical ports, the PTS 410 receives media bitstreams, e.g., audio signals, video signals, command and control data, and other data that could be text or binary. Port addresses for receiving bitstreams by the PTS 410 and the destination addresses to which the transcoded streams are transmitted are specified by a messaging protocol between the call signaling gateway and the PTS 410 through the NGCI 420. The underlying physical interface of the MNI 430 could be a Gigabit network interface card, time domain multiplexed (TDM) circuit-switched connections such as E1/T1/OC3. The MNI 430 physical network interfaces may or may not share the physical link of the NGCI 420.

[64] A monitor and setup interface (MSI) 440 is used for initial configuration, monitoring, and reconfiguration of the PTS 410. The physical link over which the MSI data are carried may be shared with other network connections or may be carried through a dedicated serial connection port. Many connectivity scenarios exist for a PTS, including connecting to a gateway e.g., a media gateway controller, and connecting to a content server, e.g., a video server.

[65] Figure 5 illustrates the connectivity of a PTS according to an embodiment of the present invention. This diagram is merely an example, which should not unduly limit the scope of the claims herein. One of ordinary skill in the art would recognize many other variations, modifications, and alternatives. A MNI 530 connects a PTS 510 to the Internet 550, via a router or switch 540, and a NGCI connects the PTS 510 to a content server (or gateway 560 can be a gateway mediating between two end-points, e.g. a mobile and an IP phone) 560 via the NGCI 520. The MSI 530 may also in connection to the content server 560 if the content server can provide basic terminal emulation support, e.g., hyper-terms under Windows. In one embodiment, the gateway or content server may be a network host.

[66] Figure 6 illustrates various PTS functions the means of an operation cycle according to an embodiment of the present invention. This diagram is merely an example, which should not unduly limit the scope of the claims herein. One of ordinary skill in the art would recognize many other variations, modifications, and alternatives. Once a PTS is connected to the network, e.g., a network host or a router, and is powered-up and configured, the PTS is ready to accept the initiation of transcoding sessions from the network host.

WO 2002/073443

PCT/US2002/069218

[67] At step 606, the PTS receives a message from either a gateway or from a content server. At step 608, the PTS determines if the received message is a session message. If the message is a session message, at step 632, the PTS decides if the message is a start of session message. A session may be initiated by a network host through a messaging protocol operated over the NGCI, such the MGCP or the H.248/MEGACO protocols. One or more sessions can be active at any time. Once a session is started, basic operations of the PTS for that session are controlled by further messages.

[68] If the message received is a start of session message, at step 650, the PTS starts a new session. If the received message is not a start of session message, at step 634, the PTS decides whether the received message is a session maintenance message. If it is a session maintenance message, the PTS commences a session maintenance process at step 652. The session maintenance process may involve maintaining or terminating a connection for a session. If the received message is not a session maintenance message, at step 636, the PTS decides whether the received message is a transcoding message. If the message is a transcoding message, at step 654, the PTS processes the transcoding message. If the message is not a transcoding message, at step 638, the PTS determines whether the received message is a rate control message. If the received message is a rate control message, at step 656, the rate control message is processed. The rate control process may direct the PTS to dynamically adjust the transmission rate in order to efficiently utilize the transmission bandwidth and to prevent transmitted packets from being dropped.

[69] When the PTS transcodes a bitstream, the optimal data rate produced by the PTS depends on media protocol of the network between the PTS and the destination of the bitstream. The data rate received by the destination varies, depending on network congestion, type of link between a router and the destination, e.g., wired connection or wireless connection, protocol and data multiplexing associated to the link, and the quality of the link. The destination may receive a data rate higher than it can process, resulting in buffer-overflow. On the other hand, the destination may receive a data rate lower than the rate it expects to receive, resulting in buffer-underflow. Therefore, the PTS may have to regulate its data rate to avoid both buffer-overflow and buffer-underflow.

[70] The PTS regulates its data rate with at least the following methods. The PTS may use network congestion information, bandwidth information, quality information from

WO 2002/073443

PCT/US2001/006218

the network host or the network access provider, or from internal PTS mechanisms to calculate roundtrip time. The roundtrip time between the PTS and an end-point, e.g., source or destination of a bitstream, can be measured by sending a "ping" packet to the end-point. The total time for the "ping" packet to reach the end-point and for the response packet from the end-point to reach the PTS is the roundtrip time. The more congested the network is, the longer is the roundtrip time. Hence the roundtrip time can be used to assess the congestion level of the network at current PTS bit-rate.

[71] Alternately, the PTS may assess the network congestion using in-band information. For example, the PTS may receive from an end-point, instructions to reduce or to increase its bit-rate throughput, under protocols such as H.245 used in H.324 and H.323.

[72] Either the in-band method or the roundtrip-time method may be used to maintain adequate quality of service given instantaneous network conditions. The PTS uses the congestion and bit-rate information to determine an adequate coding mode that yields an appropriate bit-rate with the following methods.

[73] In one embodiment, the PTS may change transcoding parameters to satisfy service goals in real-time operation. For example, in MPEG4-video, the quantization parameters can be changed to yield a low bit-rate throughput. But the change may lead to degradation in video quality. Therefore if video quality is important, advanced coding techniques such as that provided by H.263 and MPEG4 may be used to reduce the bit-rate throughput without degrading quality. But these methods may impose higher computation requirements. Hence, the PTS rate control strategy should provide a desired balance between signal quality, bit-rate and computation.

[74] The rate control in audio is achieved similarly although most audio codecs cannot provide fine-grained variable rates, but instead provide a number of bit rates that the PTS can select from. For example, the G.723.1 audio codec provides two bit-rates, a low rate and a high rate. Similarly, the GSM-AMR codec supports eight bit-rates, ranging from 4.75 Kbps to 12.2 Kbps. The PTS may use a lower rate if the network path to the end-point has congestions, or if the bandwidth allocated to the link with the end-point has a low bandwidth.

[75] In yet another embodiment, a PTS may instruct the network equipment such as a router to give a higher priority to the data being handled by the PTS, if the network equipment supports prioritization of data delivery. For instance, version 6 of the Internet

WQ 2002/073443

PCT/US2002/008218

protocol provides support for prioritization of packets. Also, the IETF has developed standards for resource reservation that allows end-points to reserve bandwidth. The PTS may exploit this packet prioritization and resource reservation if they are supported by network in which the PTS is being deployed. For example, the Internet protocol provides facility for assigning priority to packets, and the PTS can use this facility to prioritize packets when needed. In addition, the PTS may support internal prioritization of its own connections so connections of higher priority are processed as soon as they become practical.

[76] Turning back to Figure 6, at step 640, the PTS determines whether the message received is a capability message. A capability message is a message that contains capabilities of the end-points. If the message is a capability message, at step 658, the PTS processes this message. The PTS can process the message defining capabilities of end-points. For example, if the end-points are a video server and a mobile terminal, the capabilities of the video server and the mobile terminal are conveyed to the PTS. Consequently, the PTS determines the best mode of communication between these two end-points. In choosing the best mode, the PTS takes into account the protocol associated with the specific video content that the mobile user would like to view. Different types of contents may be encoded with different protocols, e.g., MPEG2 and MPEG4. In addition, the capabilities of the mobile terminal may be communicated to the PTS via the network host in numerous ways. For example, the PTS may obtain the capabilities from information stored in the mobile terminal, from the user subscription information stored in the network database of the user's service provider, or in-band within the bitstream exchanged between the mobile terminal and the network access gateway during the call signaling phase. The format of the capabilities includes among others ITU, IETF, and WAP.

[77] A capability message may be sent by the network host to the PTS to find the best transcoding mode for a specific media to be transmitted from one end-point, e.g., a video server, to the other end-point. In the capability-mode-selection process, the PTS may choose one bitstream protocol mode for receiving data from the source, and another bitstream protocol mode to which the PTS convert the received media.

[78] In yet another embodiment, the selected modes for each end-point may be signaled to the respective end-point in order to open bitstream transport channels. In H.323 or H.324, the network uses H.245 logical channel operations or Fast Connect procedures to open such channels. Using H.245 logical channel operations, an end-point may send an "open logical

WO 2002/073443

PCT/US2002/008218

channel" request to the other end-point in order to transmit signals. In H.323, an end-point can encapsulate information about media channels through which the end-point is ready to transmit signals by using a "fast start" message encapsulated in the call signaling information, exchanged in the initial setup of the call under, for example, the ITU Q.931 standard as recommended by H.225.0. Therefore, given the selected protocols for the end-points, the network host or the PTS can establish transport channels for a media bitstream. The opening process depends on the system level protocol of the overall connection between the end-points mediated by the network.

[79] Before the PTS can transcode, the selected media transmission modes have to be associated with a source address and a destination address, and the information of such association has to be communicated to the PTS via messages. The association may result from opening a logical media channel by one of the end-points as done in the H.245 standard. This implicit opening of a logical media channel when the PTS selects the types of media transmission modes can be explicitly requested by the gateway or the content server, or could be pre-programmed under some standard requirements. Regardless of the means for specifying the source address or the destination address, the association between selected media transmission modes and the source or destination addresses informs the PTS where to get the input bitstream and where to send the transcoded bitstream.

[80] In a specific embodiment, the PTS reads the bitstream from the source address, translates the bitstream from its original format to the target format, and sends the converted bitstream to the destination address. The reception and transmission of bitstream data is performed with network read/write functions, using the network hardware-specific software.

[81] Turning back to Figure 6, at step 642, the PTS determines whether the message received is a network-addressing message. A network-addressing message contains information about network addresses for the source and/or the destination of a bitstream. If the message received is a network-addressing message, at step 660, the PTS processes the network-addressing message. At step 644, the PTS determines whether the message received is a media-mixing message. A media-mixing message is a message requesting the PTS to mix signals associated with two or more audio streams and retransmit the mixed bitstreams to a network destination address. If the received message is a media-mixing message, at step 662, the PTS processes the media-mixing message. At step 646, the PTS determines whether the message received is a IP rights message, and if so at step 664 the PTS manages the information regarding

WO 2002/073443

PCT/US2002/008218

IP rights contained in the received message, according to instructions in the message. Some media communication and representation protocols support interfacing with IP rights management and processing, and accessing information regarding IP rights contained in bitstreams. For example, the PTS may support the MPEG-4 interface specification for IP rights management and processing. The information regarding IP rights is extracted or demultiplexed from MPEG4 bitstreams and made available to the network host via the messaging system. The IP rights specific applications installed on the network host or plugged into the PTS then access and process such information for various purposes, including record keeping, resigning of the context, and blocking.

[82] At step 610 in Figure 6, the PTS determines whether the received message is a features and mode message and if so, at step 622, the PTS processes this message. Processing a features and mode message includes activating features and options that are related to transcoding, mixing and other options associated to the session. At step 612, the PTS determines whether the received message is a resource message, and if so, at step 624, the PTS processes this message. Processing a resource message includes processing the PTS resources in terms of hardware processing resources, memory resources, and other computing or networking resources that the PTS manages. If the PTS receives instructions that the media being received and thereafter transmitted needs to be encrypted, the PTS transcodes and thereafter encrypts the transmitted data. Similarly, the PTS may decrypt data in response to instructions received.

[83] In the above, the capabilities of the PTS are illustrated as discrete steps, but one skilled in the art will appreciate that one or more of the discrete steps may be combined or further subdivided to perform functions of the discrete steps. Depending upon the embodiment, the functionality described can be separated or even combined. The functionality can be implemented in software and/or hardware including any combination of them. Depending upon the embodiment, there can be many other modifications, variations, and alternatives.

[84] Figures 7, 8, 9, and 10 are simplified flow diagrams illustrating an embodiment of the main system messages. These diagrams are merely examples, which should not unduly limit the scope of the claims herein. One of ordinary skill in the art would recognize many other variations, modifications, and alternatives. The main system messages allows the PTS to respond to instructions from a network equipment such as a media gateway controller. The messages includes instructions to start a session, end a session, set session options, get

WO 2002/073443

PCT/US2002/009219

session options, send a message to session manager, get message from session manager, set PTS mode, get PTS mode, set PTS feature, get PTS feature, get resource status, set resource status, update firmware procedure, get PTS system status, reset the PTS, shutdown the PTS, and activating debug/tracing mode.

5 [85] Figures 11-16 are simplified flow diagrams illustrating an embodiment of the PTS session maintenance and transcoding messages. These diagram are merely examples, which should not unduly limit the scope of the claims herein. One of ordinary skill in the art would recognize many other variations, modifications, and alternatives. The session maintenance and transcoding messages allow the PTS to terminate a session, open/close transcoding channel, 10 set/get transcoding options, updating end-point capabilities, match end-points capabilities, select end-point capabilities, activate rate control, get rate control mode, get rate control mode, add/remove media destination addresses, add/remove media source addresses, set/get IPR modes, set IPR options, activate mixing of media channels, disable mixing, and set channel mixing mode.

15 [86] Figure 17 shows a block diagram illustrating software modules that may be used in a PTS according to an embodiment of the present invention. This diagram is merely an example, which should not unduly limit the scope of the claims herein. One of ordinary skill in the art would recognize many other variations, modifications, and alternatives. The PTS software comprises the following main modules:

- 20 [87] 1. A session management module 1710;  
 [88] 2. A PTS management module 1720;  
 [89] 3. A network host interface module 1730;  
 [90] 4. A media channel processing module 1740;  
 [91] 5. A call signaling interface module 1750;  
 25 [92] 6. A network interface module 1760;  
 [93] 7. A transcoding module 1770;  
 [94] 8. A rate control module 1780;  
 [95] 9. An intellectual property rights management module 1790; and  
 [96] 10. A capabilities processing module 1792.  
 30 [97] The session management module 1710 performs the main services of a gateway, and is thus the main PTS software program. For example, the session management



WO 2002/073443

PCT/US2002/008218

module 1710 starts and ends transcoding sessions, handles and dispatches session messages, and manages session resources. The PTS management module 1720 performs basic overall management functions necessary if an operator needs to inspect the status of the PTS or to manage its resources. For example, the management module 1720 tests main hardware components with component specific testing procedures, resets the PTS, and tracks and allocates transcoding resources dynamically. The network host interface module 1730 handles communication-messaging interface between the PTS and the network host, e.g., a media gateway controller or a content server. For example, the interface module 1730, depending on type of network host, implements messaging between the network host and the PTS, implements commands defining methods by which transcoding capabilities are captured by the PTS or defined for the PTS. In addition, the interface module 1730 may implement methods by which the PTS retrieves media content types from bitstreams when capability exchanges are not explicitly performed. The media channel processing module 1740 performs media channel function, such as opening, closing, adding and removing channel network sources and destinations. The call signaling interface module 1750 performs functions for establishing initial call setup between end-points through the PTS, where the procedure for call setup depends on standard, such as SIP and Q.931. The network interface module 1760 provides basic input and/or output communication interfaces. The basic input and/or output is the lowest level of communication over which more complex messaging is performed.

[98] The transcoding module 1770 performs the actual transcoding function including transcoding between MPEG series, H.26X video series, GSM-AMR and G.72X audio codec series. Other examples of PTS transcoding may include at least MPEG2-audio to MPEG4-audio, G.723.1 to GSM-AMR, MPEG2-video to MPEG4-video, H.263 to MPEG4-video. The rate control module 1780 performs a throttling function and prevents network segments over which the end-points communicate from being overloaded. The intellectual property rights management module 1790 protects the IP rights. For example, the management module 1790 uses data on IP rights in a bitstream, and assists on implementing mechanisms for auditing, monitoring, billing, and protection of the IP rights associated with the received and transcoded bitstream. The capabilities processing module 1792 utilizes data in received messages to find the best matching mode for the specific media that is transmitted from one end-point to another.

WO 2002/073443

PCT/US2002/005218

[199] The architecture of a PTS determines performance, cost, and time-to-market of the server. Performance may be considered as the number of simultaneous gateway channels or calls that the PTS may simultaneously process. Furthermore, for a fixed number of channels, the cost and performance of the architecture will depend on the following factors:

- 5 [100] 1. Bus architecture;
- [101] 2. Transcoding architecture and hardware for various video and audio transcoding;
- [102] 3. Network off loading for connecting to MGC and other gateway components; and
- 10 [103] 4. Operating System.

[104] Figure 18 shows an explanation of the symbols used in the flow charts. This diagram is merely an example, which should not unduly limit the scope of the claims herein. One of ordinary skill in the art would recognize many other variations, modifications, and alternatives.

- 15 [105] Figure 19 shows a simplified flow chart illustrating the high-level procedure for video bitstream transcoding that may be used in a PTS according to an embodiment of the present invention. This diagram is merely an example, which should not unduly limit the scope of the claims herein. One of ordinary skill in the art would recognize many other variations, modifications, and alternatives. The procedure reads a chunk of bits and if
- 20 the end-sequence marker is detected then it ends. Otherwise it reads the next code word, transcodes the code word to the output protocol codeword, updates the history record and emits the transcoded bits to the output buffer which is flushed according to a rate control scheme as to avoid that the input buffer at the receiving end-point does not over-flow.

- [106] Figure 20 illustrates a PTS hardware architecture according to an
- 25 embodiment of the present invention. This diagram is merely an example, which should not unduly limit the scope of the claims herein. One of ordinary skill in the art would recognize many other variations, modifications, and alternatives. The architecture features an intelligent transcoding node 2010, which is a bus card including one or more processors, a bank of DSP processors 2020, one or more network interfaces 2030, one or more processors 2040, and a
- 30 memory bank 2050. This architecture has several advantages. First, the network interface 2030 is embedded in the intelligent transcoding node bus-card 2010. Hence the call processing and

WG 2002073443

PCT/US2002/008218

transcoding are performed locally on the bus-card 2010. Second, one or more network interfaces per processing node are possible. Third, the architecture can support a large number of simultaneous calls because of compactness of its processing modules.

[107] In addition to the embodiment in Figure 20, numerous architectures for a

5 PTS are possible, a few of which are listed below:

- [108] 1. Standalone chassis with bus-cards;
- [109] 2. PC-like implementation as described below;
- [110] 3. Firmware to existing processing hardware, including ASICs;
- [111] 4. Software running on existing hardware;
- 10 [112] 5. Software running on existing hardware with hardware acceleration by the means of ASICs, DSPs, or other types of processors; and
- [113] 6. ASIC chipset.

[114] Figure 21 shows an embodiment of a computer system according to the present invention. This diagram is merely an example, which should not unduly limit the scope of the claims herein. One of ordinary skill in the art would recognize many other variations, modifications, and alternatives. The present invention may be implemented in a personal computer (PC) architecture. Alternative computer system architectures, or other programmable or electronic-based devices may also be employed.

[115] In Figure 21 a computer system 2100 comprises a bus 2101 for  
 20 communicating information, a processor 2102 coupled to the bus 2101 for processing information, random access memory 2103 coupled to the bus 2101 for storing information and instructions for the processor 2102, a read-only memory 2104 coupled to the bus 2101 for storing static information and instructions for the processor 2102 and PTS applications, a display device 2105 coupled to the bus 2101 for displaying information for a user, an input device 2106 coupled  
 25 to the bus 2101 for communicating information and command selections to the processor 2102, and a mass storage device 2107, e.g., a magnetic disk and associated disk drive, coupled to the bus 2101 for storing information and instructions. A data storage medium 2108 contains digital information, e.g., PTS software modules, and it is coupled to the bus 2101 and configured to operate with mass storage device 2107 to provide the processor 2102 access to digital  
 30 information stored on the data storage medium 2108 through the bus 2101. A hardware transcoding acceleration module 2109 comprises printed circuit boards, Digital Signal

WO 2002/073443

PCT/US2002/009218

Processors, ASIC's, and FPGA's. The module 2109 is communicatively coupled to the bus 2101, and may have an embodiment similar to the intelligent transcoding node 2010 as shown in Figure 20.

[116] Processor 2102 may be any of a wide variety of general purpose  
 5 processors or microprocessors, e.g., the Pentium<sup>TM</sup> processor manufactured by Intel Corporation, and a MIPS processor manufactured by MIPS Technologies, Inc., of 2011 N. Shoreline Blvd., Mountain View, CA 94039-7311. Other varieties of processors such as digital signal processors (DSP's) may also be used in the computer system 2100. The display device 2105 may be a liquid crystal device, cathode ray tube (CRT), or other suitable display device. The mass storage  
 10 device 2107 may be a conventional hard disk drive, a floppy disk drive, a CD-ROM drive, or other magnetic or optical data storage device for reading and writing information stored on a hard disk, a floppy disk, a CD-ROM, a magnetic tape, or other magnetic or optical data storage medium. The data storage medium 2108 may be a hard disk, a floppy disk, a CD-ROM, a magnetic tape, or other magnetic or optical data storage medium.

[117] In general, the processor 2102 can retrieve processing instructions and  
 15 data from the read-only memory 2104. The processor 2102 can also retrieve processing instructions and data from the data storage medium 2108 using the mass storage device 2107 and downloads the information into the random access memory 2103, which may be a SDRAM. The processor 2102 then executes an instruction stream from the random access memory 2103 or  
 20 read-only memory 2104. Command selections and information input at the input device 2106 may direct the flow of instructions executed by the processor 2102. Input device 2106 may be, among others, a pointing device such as a conventional mouse or a trackball device. The execution results may be displayed on the display device 2105. The computer system 2100 also comprises a network device 2110 for connecting the computer system 2100 to a network. The  
 25 network device 2110 may be an Ethernet device, a phone jack, a satellite link, or other device.

[118] Embodiments of the present invention may be represented as a software  
 product stored on a machine-accessible medium, also referred to as a computer-accessible  
 medium or a processor-accessible medium. The machine-accessible medium may be any type of  
 30 magnetic, optical, or electrical storage medium including a diskette, a CD-ROM, a memory device, either volatile or non-volatile, an ASIC, a firmware to an ASIC, a system-on-chip, or other storage mechanism. The machine-accessible medium may contain various sets of

WO 2002/073443

PCT/US2002/008218

instructions, code sequences, or configuration information. Other data necessary to implement the present invention may also be stored on the machine-accessible medium. As merely an example, transcoding techniques are described in U.S. Provisional Serial No. 60/347270 (Attorney Docket No. 021318-000200US), commonly assigned, and hereby incorporated by reference for all purposes.

[119] In addition to media transcoding that the PTS can perform, it can also perform system protocol transcoding. A multimedia system protocol is typically an umbrella of protocols that define how multimedia end-points can connect to each other, can issue and interpret commands (such streaming or opening a video channel), can tear down connections, can join conferences. A system protocol typically covers the following important aspects: call signalling, command and control, media transports aspects, and media coding aspects. For example the H.323 system protocol standard covers H.225.0/Q.931 for call signalling and media transports, H.245 for command and control, and a number of audio and video codecs. The H.324 system protocol standard covers H.223 (media and data bitstreams multiplexing), H.245 for command and control, and a number of audio and video codecs. Although some aspects of H.323 and H.324 are similar, H.323 is packet based whereas H.324 is circuit based. Hence for H.323 and H.324 end-points to communicate, system protocol transcoding needs to be performed, at the call-signalling level, command and control levels and media coding levels. The PTS performs both system protocol transcoding and media (audio and video) transcoding. From the PTS point of view, the Call signalling transcoding is the process of proxying the end-points (H.324, H.323, SIP, RTSP, and others) so they can connect to each other. From the command and control point of view, the PTS performs the transcoding so messages such as terminal capabilities, open/close logical channels, etc., are translated so they can be understood by the terminal receiving the command and control message. The PTS has to perform translation of the issued messages it receives from the sending end-point so they could be understood by the receiving terminal. In terms of media transport, the PTS needs to access data by demultiplexing from circuit bearer channels, extracting the media service data units, transcoding the media bits and then packaging the transcoded bits so they can be sent in a format the receiving end point can understand. In the receiving end point is H.323, the packaging would involved RTP packetization. So the media transport transcoding (translation) consists of circuit-to-packet, packet-to-circuit and packet-to-packet translation of the media bits. When the transport is circuit

WO 2002/073443

PCT/US2002/008218

based, the typically the media bits are multiplexed in a Time Domain Multiplexing (TDM) fashion.

[120] While there has been illustrated and described what are presently considered to be example embodiments of the present invention, it will be understood by those skilled in the art that various other modifications may be made, and equivalents may be substituted, without departing from the true scope of the invention. Additionally, many modifications may be made to adapt a particular situation to the teachings of the present invention without departing from the central inventive concept described herein.

WO 2002/073443

PCT/US2002/009219

WHAT IS CLAIMED IS:

- 1           1.     A system for transferring multimedia information from a source location  
2     to a destination location through one or more networks, the system comprising:  
3         a source output providing a first stream of information in one of a plurality of  
4     source;  
5         a destination input receiving a second stream of information in one of a plurality  
6     of destination capabilities;  
7         a proxy transcoder server ("PTS") coupled between the source output and the  
8     destination input, the PTS comprising:  
9         a capability module adapted to identify the source capability of the source  
10     output and adapted to identify destination capability of the destination input;  
11         a selection module adapted to select a transcoding process based upon the  
12     one capability of the source capabilities and the one capability from the destination  
13     capabilities;  
14         a transcoding module adapted to use the selected transcoding process to  
15     process the first stream of information.
- 1           2.     The system of claim 1 wherein the one or more transport networks are  
2     selected from a group comprising the Internet, a mobile network, a wide area network, a local area  
3     network, PTN, ISDN, and SONET.
- 1           3.     The system of claim 1 wherein at least one of the source output and the  
2     destination input is that of a remote device.
- 1           4.     The system of claim 3 wherein the capability module identifies at least one  
2     of the output and input of the remote device, based on information stored in the device, based on  
3     user subscription information stored in a network database of the user's service provider, based  
4     on in-band information command and control within a stream exchanged, or pre-set by the  
5     service provider.

WO 2002/073443

PCT/US2002/008218

- 1           5.     The system of claim 1 wherein the transcoding process selected by the  
2 capability module transcodes data from a first bitstream protocol mode to a second bitstream  
3 protocol mode.
- 1           6.     The system for claim 1 wherein the PTS further comprising a rate control  
2 module regulating the data rate produced by the PTS.
- 1           7.     The system for claim 6 wherein the rate control module detects network  
2 status information by calculating "round-trip" time information based on network congestion  
3 information, bandwidth information, quality information from a network host or network access  
4 provider, or internal PTS mechanisms.
- 1           8.     The system for claim 7 wherein the "round-trip" time information can be  
2 measured by send a "ping" packet to either the source location or the destination location.
- 1           9.     The system for claim 6 wherein the rate control module detects the  
2 network status information by using in-band information.
- 1           10.    The system for claim 6 wherein the rate control module regulates the data  
2 rate by changing transcoding parameters.
- 1           11.    The system for claim 6 wherein the rate control module regulates the data  
2 rate by instructing network equipment to give a higher priority to data being handled by the PTS  
3 than other data.
- 1           12.    The system of claim 1 wherein the format of the capability is selected  
2 from a group comprising ITU, IETF, and WAP.
- 1           13.    The system of claim 1 wherein the one or more networks are selected from  
2 a plurality of different networks, each of the network being configured for a particular standard.
- 1           14.    The system of claim 1 wherein the PTS further comprising a network  
2 addressing module to determine the network address of the source output and the network  
3 address of the destination input.



WO 2002/073443

PCT/US2002/008218

1           15. The system of claim 1 wherein the PTS further comprising a media mixing  
2 process to combine bitstreams associated with two or more audio streams and retransmit the  
3 combined bitstreams to the destination input.

1           16. The system of claim 1 wherein the PTS further comprising an intellectual  
2 property rights management module to manage and process information on intellectual property  
3 rights.

1           17. The system of claim 1 wherein the PTS further comprising a encryption  
2 and decryption process to encrypt and decrypt the data.

1           18. The system for claim 6 wherein the rate control module regulates the data  
2 rate dynamically and in real time.

1           19. The system of claim 1 wherein the transcoding module are programmable  
2 to transcode between various types of capabilities for the source output and various types of  
3 capabilities for the destination input.

1           20. A system for transferring multimedia information from source to  
2 destination locations through one or more networks, the system comprising:  
3 a source output in a first format from a plurality of source capabilities, the source  
4 output being coupled to a first network, the source output providing a first stream of information;  
5 a destination input to be received in a second format from a plurality of  
6 destination capabilities, the destination input being coupled to a second network, the destination  
7 input receiving a second stream of information;  
8 a proxy transcoder server ("PTS") coupled between the source output and the  
9 destination input, the proxy transcoder server comprising:  
10 a capability process coupled to the source output, the capability process  
11 being adapted to identify the first format of the source output and adapted to identify the second  
12 format of the destination input;  
13 a transcoding process coupled to the capability process, the transcoding  
14 process comprising a plurality of transcoding modules numbered 1 through N, where N is an  
15 integer greater than 1, the transcoding process being adapted to select one of the transcoding

WO 2002/073443

PCT/US2002/008218

16 process based upon the first format that is associated with a capability and the second format that  
17 is associated with a second capability; and  
18 a bit rate control process coupled to the transcoding process, the bit rate  
19 control process being adapted to receive a network status information from the first network, the  
20 bit rate control being adapted to adjust a status of the stream of information based upon the  
21 network status information.

1 21. The system of claim 20 wherein the status information comprises a ping.

1 22. The system of claim 20 wherein the status is a stop status.

1 23. The system of claim 20 wherein the status is a prioritization status.

1 24. The system of claim 20 wherein the status is to adjust a bit rate by  
2 selecting a lower bit rate coder.

1 25. A method for processing streams of information, the method comprising:  
2 identifying a source capability from a plurality of source capabilities for a stream  
3 of information;  
4 identifying a destination capability from a plurality of destination capabilities;  
5 selecting a transcoding process from a plurality of transcoding processes in a  
6 library based upon the identified source capability and the identified destination capability;  
7 processing the stream of information using the selected transcoding process if the  
8 identified source capability and the identified destination capability are different;  
9 transferring the stream of information from the source to the destination free from  
10 one of the transcoding processes of the identified source capability and the identified destination  
11 capability matches.

1 26. The method of claim 25 wherein the selected transcoding process is  
2 provided by empirical information.

1 27. The method of claim 25 wherein the library is a look up table having at  
2 least the plurality of source capabilities and the plurality of destination capabilities in a second  
3 dimension.

WO 2002/073443

PCT/US2002/009218

1/18

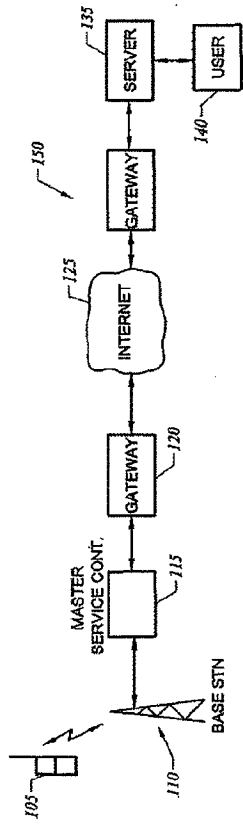


FIG. 1  
(Prior Art)

WQ 2002/073443

PCT/US2002/009218

2/18

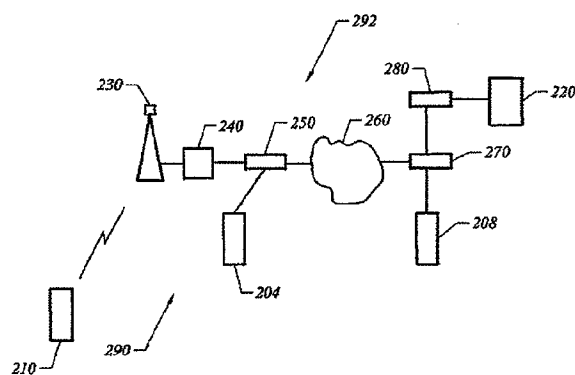


FIG. 2

SUBSTITUTE SHEET (RULE 26)

WO 2002/073443

PCT/US2002/009218

3/18

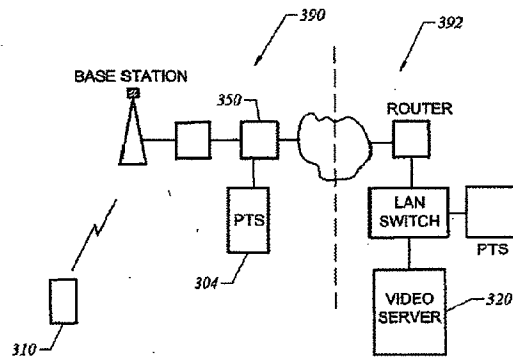


FIG. 3

WO 2002/073443

PCT/US2002/009218

4/18

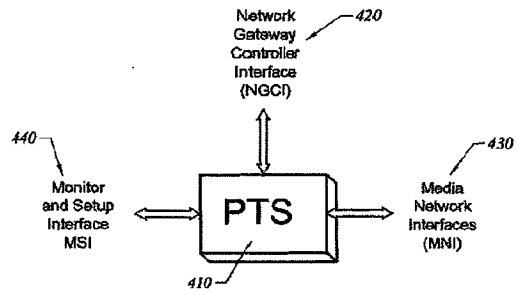


FIG. 4

W/O 2902/073443

PCT/US2002/008218

5/18

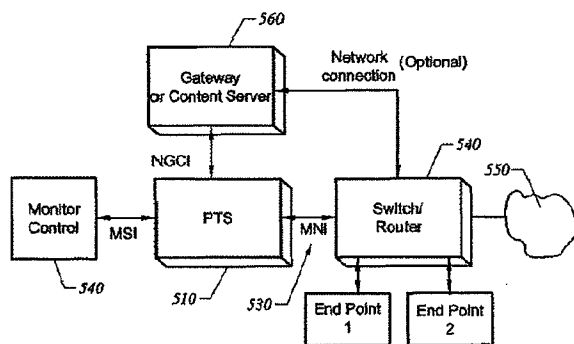


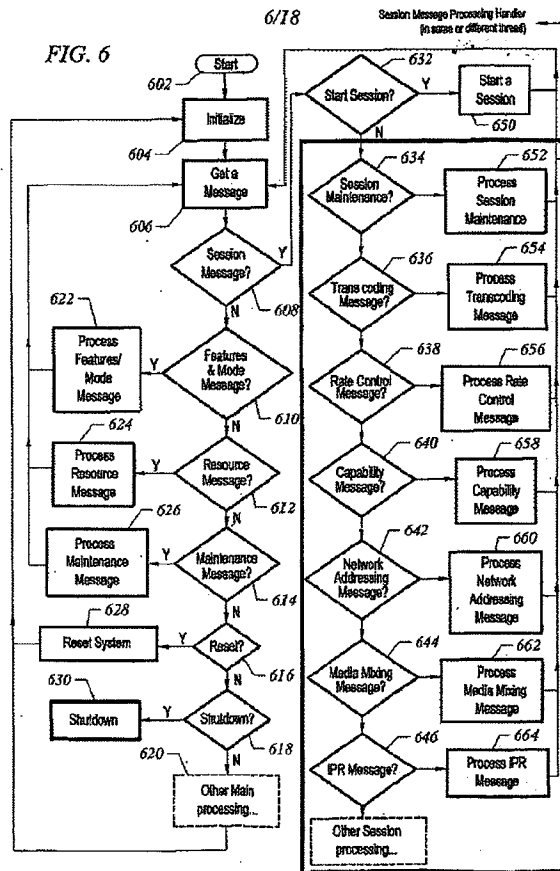
FIG. 5

SUBSTITUTE SHEET (RULE 26)

WO 2002/073443

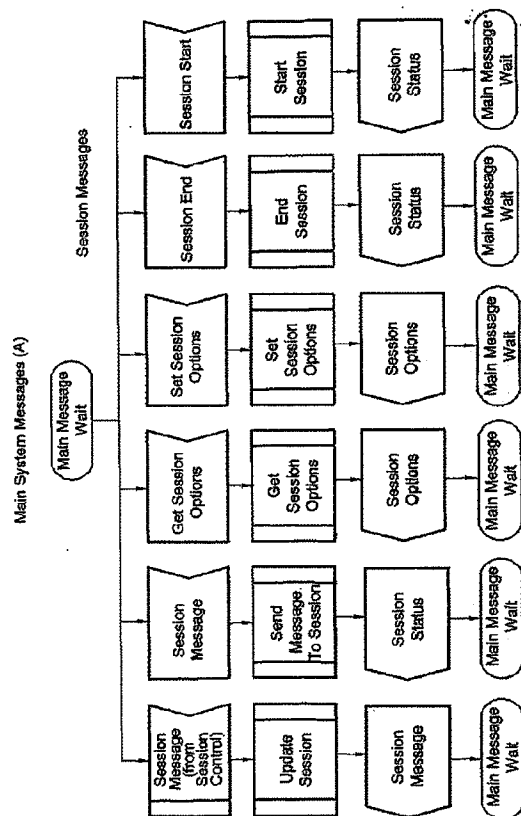
PCT/US2002/008218

FIG. 6



SUBSTITUTE SHEET (RULE 26)





WO 2002/073443

PCT/US2002/009219

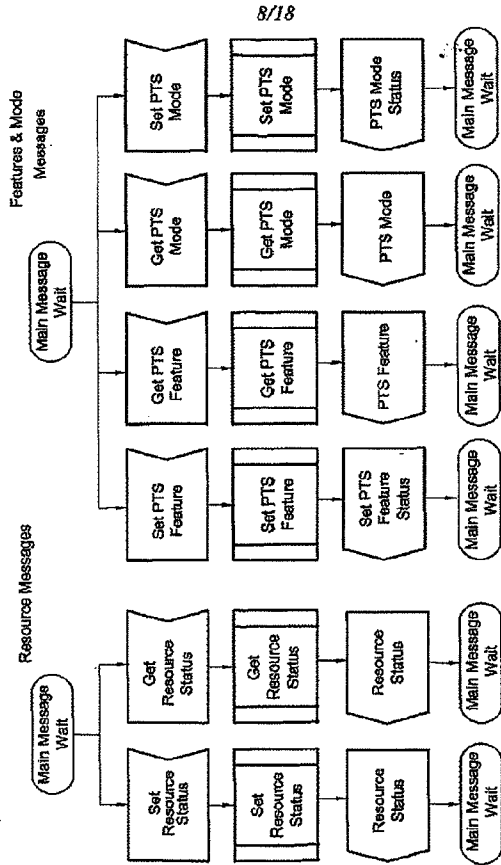


FIG. 9

FIG. 8

WO 2002/073443

PCT/US2002/008213

9/18

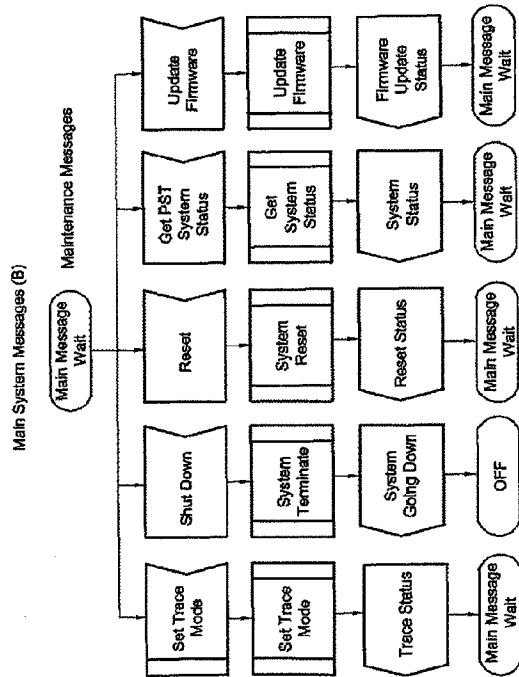
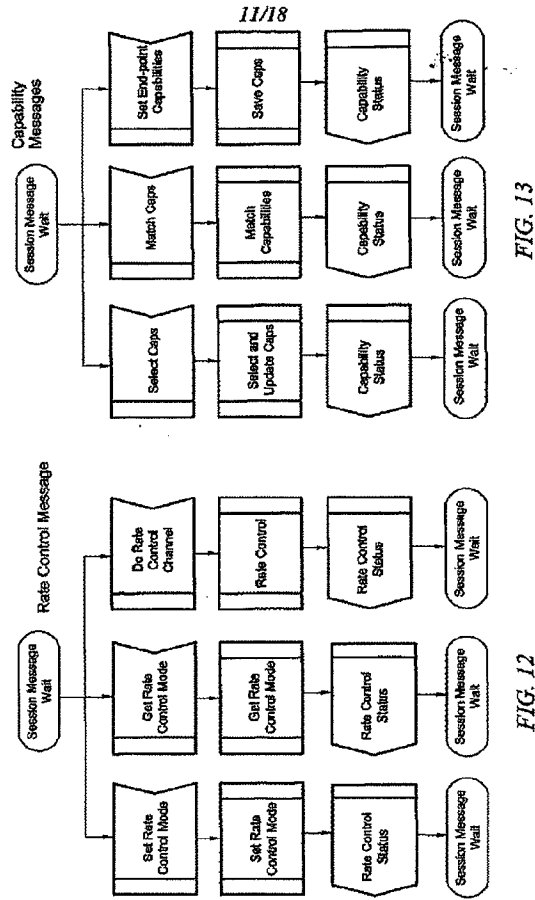


FIG. 10

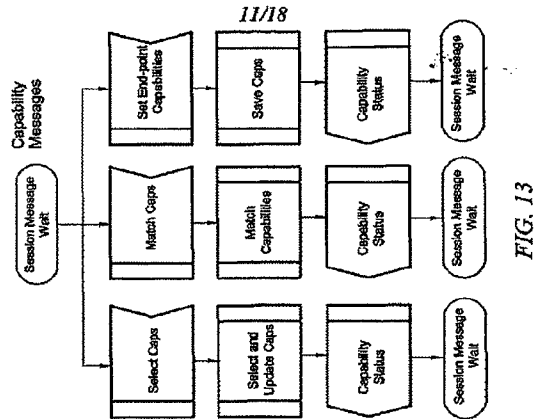


WO 2002/073443

PCT/US2002/008218



SUBSTITUTE SHEET (RULE 26)



WO 2002/073443

PCT/US2002/098218

12/18

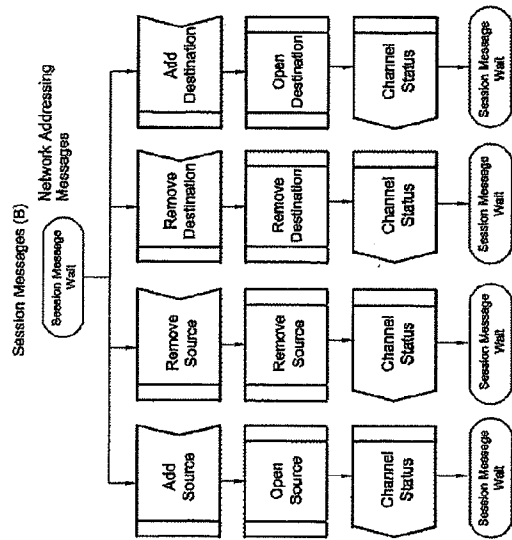


FIG. 14

WO 2002/073443

PC3/US2002/008219

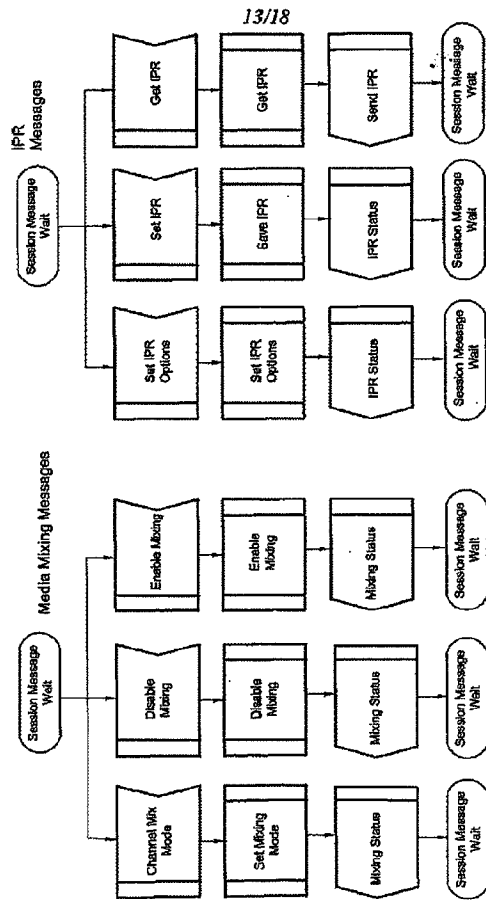


FIG. 15

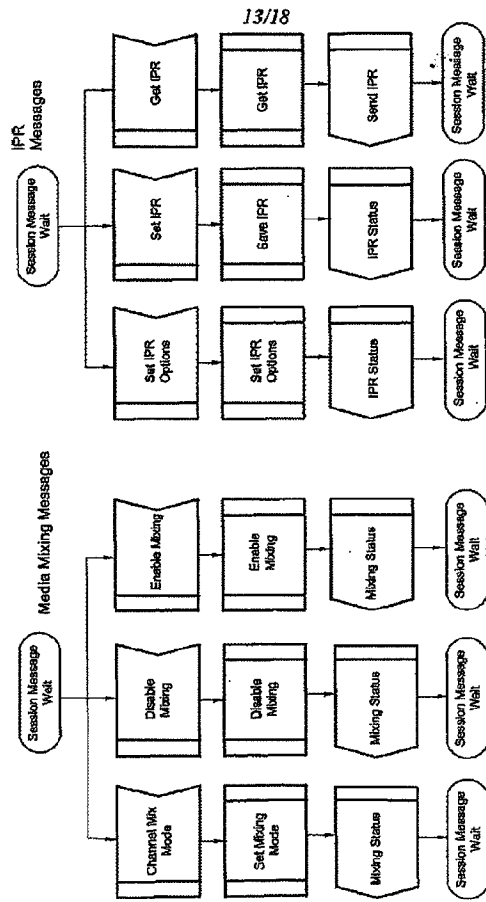


FIG. 16

WO 2002/073443

PCT/US2002/008218

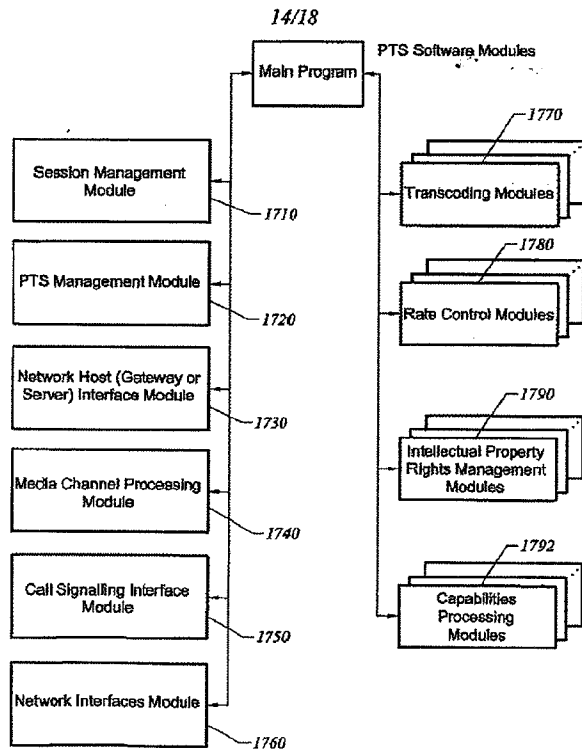


FIG. 17

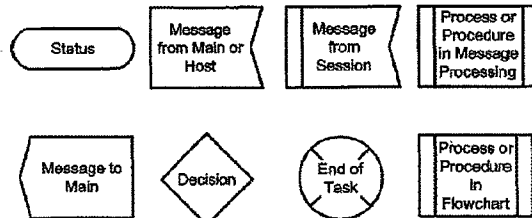


WO 2002/073443

PCT/US2002/009218

15/18

## Symbols



Symbols used in flowcharts and message diagrams

FIG. 18

WQ 2002/073443

PCT/US2002/008218

16/18

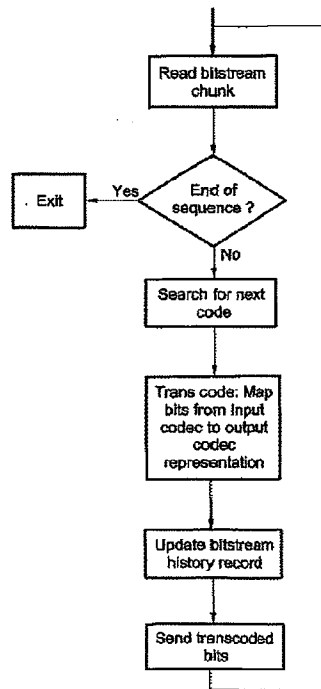


FIG. 19

WO 2002/073443

PCT/US2002/009218

17/18

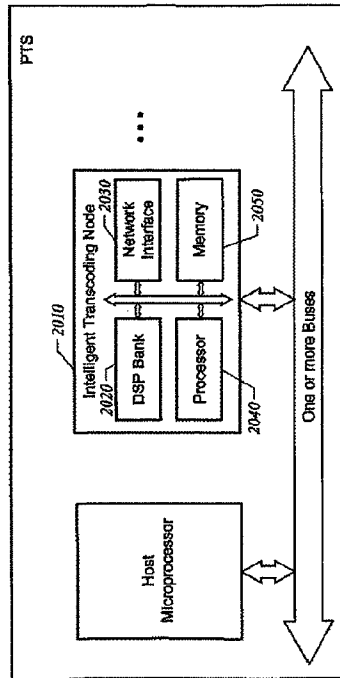


FIG. 20

SUBSTITUTE SHEET (RULE 26)

18/18

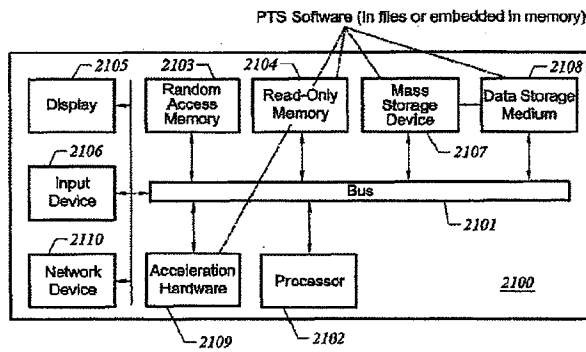


FIG. 21

## 【国際調査報告】

INTERNATIONAL SEARCH REPORT		International application No. PCT/US2002/015
<b>A. CLASSIFICATION OF SUBJECT MATTER</b> IPC Class. G06F 15/16; H04N 7/00 US CL. 705/204, 201 According to International Patent Classification (IPC) or in both national classification and IPC		
<b>B. FIELDS SEARCHED</b> Minimum documentation searched (classification system followed by classification symbols) U.S. 705/204, 201 Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) USPAT FAST Text search engine, encoding, decoding, transcribing, proxy, client-server, user/client capability, creating, user/client profile/information, internet		
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 5,940,391 A (MALKIN et al.) 17 AUGUST 1999 Abstract, Figures 1, 4A, 4B, 12. Column 1, Line 60 through Column 2, Line 2, Column 2, Line 64 through Column 8, Line 21	1-27
Y	US 5,996,022 A (KRUEGER et al.) 20 NOVEMBER 1999 Abstract, Figures 1, 4, 5, Column 2, Lines 11-21, Column 2, Line 59 through Column 3, Line 4, Column 4, Lines 36-47, Column 5, Lines 12-29, Column 5, Line 59 through Column 6, Line 67	1-27
Y, P	US 6,247,050 B1 (TSO et al.) 12 JUNE 2001 Abstract, Figures 1-7, Column 2, Lines 1-19, Column 2, Line 47 through Column 4, Lines 29, Column 5, Lines 41-54, Column 7, Lines 61 through Column 8, Line 27	1-27
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C <input type="checkbox"/> See parent family annex		
* Special categories of cited documents: "X" document defining the general state of the art which is not considered to be of particular relevance "Y" earlier document published on or after the international filing date which is cited to establish the publication date of another citation or other special reason (to be specified) "Y, P" document referred to in an oral disclosure, press publication or other source "Z" document published prior to the international filing date but later than the priority date claimed "A" document published before the international filing date but not used to establish the priority date of the invention "E" document of particular relevance: the claimed invention cannot be considered to be prior art because other cited documents, such as citations being additions to a primary citation or the art. "F" document of the same patent family		
Date of the actual completion of the international search		Date of mailing of the international search report
20 MAY 2002		26 JUN 2002
Name and mailing address of the ISA/US Executive Director of Patents and Trademarks Box PCT Washington, D.C. 20501 Facsimile No. (703) 305-3230		Authorized officer MARC THOMPSON Telephone No. (703) 305-3230 <i>R. Mehta</i>

Form PCT/ISA/210 (second sheet) (July 1999)

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/JP 2004/000100

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y, P	US 6,345,303 B1 (KNAUERHASE et al.) 05 FEBRUARY 2002 Abstract, Figures 1-5, Column 1, Lines 44-63, Column 2, Line 44 through Column 3, Line 48, Column 6, Line 62 through Column 7, Line 46	1-27
A, P	US 6,308,222 B1 (KRUEGER et al.) 23 OCTOBER 2001 Entire document	1-27

Form PCT/ISA/210 (continuation of second sheet) (July 2002)

## フロントページの続き

(81)指定国 AP(GH,GM,KE,LS,MW,MZ,SD,SL,SZ,TZ,UG,ZM,ZW),EA(AM,AZ,BY,KG,KZ,MD,RU,TJ,TM),EP(AT, BE,CH,CY,DE,DK,ES,FI,FR,GB,GR,IE,IT,LU,MC,NL,PT,SE,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GQ,GW,ML,MR,NE,SN, TD,TG),AE,AG,AL,AM,AT,AU,AZ,BA,BB,BG,BR,BY,BZ,CA,CH,CN,CO,CR,CU,CZ,DE,DK,DM,DZ,EC,EE,ES,FI,GB,GD,GE, GH,GM,HR,HU,ID,IL,IN,IS,JP,KE,KG,KP,KR,KZ,LC,LK,LR,LS,LT,LU,LV,MA,MD,MG,MK,MN,MW,MX,MZ,NO,NZ,OM,PH,P L,PT,RO,RU,SD,SE,SG,SI,SK,SL,TJ,TM,TR,TT,TZ,UA,UG,UZ,VN,YU,ZA,ZM,ZW

(特許庁注：以下のものは登録商標)

イーサネット

Fターム(参考) 5K030 HA08 HB01 HB02 HB18 HC01 HD03 KA19 LB15